

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2005, 18:58:20 ; Search time 124.5 Seconds  
(without alignments)  
236.095 Million cell updates/sec

Title: US-09-831-951A-1  
Perfect score: 394  
Sequence: 1 NKEIDPIQKLVFKIREYKS.....TFPTKPEDPKFVLEKQA 76

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	394	100.0	76 4 AAB80963	Aab80963 Human Cou
2	392	99.5	84 4 AAG75820	Aag75820 Human COL
3	392	99.5	108 2 AAY48588	Aay48588 Human bre
4	392	99.5	108 3 AAG04005	Aag04005 Human sec
5	392	99.5	108 7 ADE63626	Ade63626 Human Pro
6	392	99.5	108 7 ADE63638	Ade63638 Human Pro
7	392	99.5	108 7 ADE63634	Ade63634 Human Pro
8	392	99.5	108 7 ADE63630	Ade63630 Human Pro
9	392	99.5	108 7 ADJ69447	Adj69447 Human hea
10	383	97.2	106 3 AAG01886	Aag01886 Human sec
11	325	82.5	114 2 AAW87500	Aaw87500 Human mit
12	325	82.5	114 3 AAY82605	Aay82605 Human mit
13	317	80.5	76 4 AAB80964	Aab80964 Rat Coupl
14	317	80.5	108 7 ADE63636	Ade63636 Rat Prote
15	317	80.5	108 7 ADE63632	Ade63632 Rat Prote
16	317	80.5	108 7 ADE63628	Ade63628 Rat Prote
17	317	80.5	108 7 ADE63624	Ade63624 Rat Prote
18	185	47.0	69 3 AAG01887	Aag01887 Human sec
19	163	41.4	39 4 AAB80972	Aab80972 Rat CF6 P
20	154.5	39.2	105 4 ABB60627	Abb60627 Drosophil
21	118	29.9	147 4 ABB57878	Abb57878 Drosophil
22	112	28.4	56 4 AAM14900	Aam14900 Peptide #
23	112	28.4	56 4 AAM14420	Aam14420 Peptide #
24	112	28.4	56 4 ABB33368	Abb33368 Peptide #
25	112	28.4	56 4 ABB32935	Abb32935 Peptide #

26	112	28.4	56 4 AAM26396	Aam26396 Peptide #
27	112	28.4	56 4 AAM26833	Aam26833 Peptide #
28	112	28.4	56 4 ABB27764	Abb27764 Human pep
29	112	28.4	56 4 ABB28193	Abb28193 Human pep
30	112	28.4	56 4 ABB18414	Abb18414 Protein #
31	112	28.4	56 4 ABB18827	Abb18827 Protein #
32	112	28.4	56 4 AAM66547	Aam66547 Human bon
33	112	28.4	56 4 AAM66519	Aam66519 Human bon
34	112	28.4	56 4 AAM53736	Aam53736 Human bra
35	112	28.4	56 4 AAM54153	Aam54153 Human bra
36	112	28.4	56 4 ABG47787	Abg47787 Human liv
37	112	28.4	56 4 ABG48215	Abg48215 Human liv
38	112	28.4	56 4 AAM02147	Aam02147 Peptide #
39	112	28.4	56 4 AAM01731	Aam01731 Peptide #
40	112	28.4	56 5 ABG35769	Abg35769 Human pep
41	112	28.4	56 5 ABG36199	Abg36199 Human pep
42	92	23.4	20 4 AAB80968	Aab80968 Rat CF6 P
43	92	23.4	20 4 AAB80969	Aab80969 Rat CF6 P
44	91	23.1	18 4 AAB80971	Aab80971 Human CF6
45	91	23.1	19 4 AAB80970	Aab80970 Human CF6

## ALIGNMENTS

RESULT 1  
AAB80963  
ID AAB80963 standard; protein; 76 AA.  
XX  
AC AAB80963;  
XX  
DT 08-JUN-2001 (first entry)  
XX  
DE Human Coupling factor 6 CF6.  
XX  
KW Human; coupling factor 6; CF6; antiinflammatory; antidiabetic; antilucer;  
KW cardiac; hypotensive; antiangiogenic; antiaesthetic; antihemetic;  
KW antithrombotic; proton-translocating ATPase; prostaglandin;  
KW cytoplasmic PLA<sub>2</sub>; cardiovascular disorder; diabetes;  
KW inflammatory disorder; rheumatoid arthritis; cerebrovascular disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200121205-A1.  
XX  
PD 29-MAR-2001.  
XX  
PF 03-AUG-2000; 2000WO-JP005210.  
XX  
PR 17-SEP-1999; 99JP-00264687.  
XX  
PA (SUNR ) SUNTORY LTD.  
XX  
PI Oseanal T, Magota K;  
XX  
DR WPI; 2001-266047/27.  
XX  
PT Coupling factor 6 useful in diagnosis of diseases relating to excess or  
PT lack of prostaglandin and cytoplasmic PLA<sub>2</sub> activity in blood, and in  
PT screening drugs for treating e.g. cardiovascular infection.  
XX  
PS Claim 15; Page 56; 70pp; Japanese.  
XX  
The present sequence is human coupling factor 6 (CF6) protein. CF6 is a  
subunit of a proton-translocating ATPase found in mitochondria. CF6 can  
be used in diagnosis of diseases relating to excess or lack of  
prostaglandin and cytoplasmic PLA<sub>2</sub> activity, and in screening inhibitors  
and potentiators as drugs to treat e.g. cardiovascular infection,  
hypertension, arteriosclerosis, angina pectoris, cardiac insufficiency,  
cerebrovascular disorders, hyperlipemia, diabetes, bronchitis, gastric  
ulcer, pregnant eclampsia, hemolytic uremia syndrome, thrombopenic  
pupura, inflammatory diseases like cerebral infarction, acute  
pancreatitis, asthma, ARDS (adult respiratory distress syndrome), and

CC rheumatoid arthritis  
XX Sequence 76 AA;

Query Match 100.0%; Score 394; DB 4; Length 76;  
Best Local Similarity 100.0%; Pred. No. 1.1e-39;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKELDPIQKLFVDKIREYKSKRQTSGGPVDASSEYQOELERELFKLKOMFGNADMTPT 60  
DB 1 NKELDPIQKLFVDKIREYKSKRQTSGGPVDASSEYQOELERELFKLKOMFGNADMTPT 60

QY 61 FKFEDEPKFEVLEKPOA 76  
DB 61 FKFEDEPKFEVLEKPOA 76

RESULT 2  
AAG75820  
ID AAG75820 standard; protein; 84 AA.

XX AAG75820;  
AC AAG75820;  
DT 03-SEP-2001 (first entry)

XX Human colon cancer antigen protein SEQ ID NO:6584.

DE Human colon cancer antigen protein SEQ ID NO:6584.  
KM Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW colorectal carcinoma; chromosome 10.

XX Homo sapiens.

PN WO200122920-A2.

PD 05-APR-2001.

XX 28-SEP-2000; 2000WO-US026524.

XX 29-SEP-1999; 99US-0157137P.

PR 03-NOV-1999; 99US-0163280P.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;

DR WPI; 2001-235357/24.

DR N-PSDB; AAH35225.

PS Claim 11; Page 8060-8062; 9803pp; English.  
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
PT useful for preventing, diagnosing and/or treating colorectal cancers.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the  
CC proteins are collectively known as colon cancer antigens. The colon  
CC cancer antigens have cytostatic activity and can be used in gene therapy  
CC and vaccine production. N and P may be used in the prevention, diagnosis  
CC and treatment of diseases associated with inappropriate P expression. For  
CC example, N and P may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of P by expressing inactive proteins or to  
CC supplement the patient's own production of P. Additionally, N may be used  
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids  
CC into a host cell and culturing the cell to express the proteins. N and P  
CC can be used in the prevention, diagnosis and treatment of colorectal  
CC carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent  
CC sequences used in the exemplification of the present invention. N.B.  
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at  
CC time of publication, meaning no sequences are present for SEQ ID NO:1027  
CC to 1052, 7921 and 7922  
XX Sequence 84 AA;  
SQ

Query Match 99.5%; Score 392; DB 4; Length 84;  
Best Local Similarity 98.7%; Pred. No. 2.1e-39;  
Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKELDPIQKLFVDKIREYKSKRQTSGGPVDASSEYQOELERELFKLKOMFGNADMTPT 60  
DB 9 NKELDPIQKLFVDKIREYKSKRQTSGGPVDASSEYQOELERELFKLKOMFGNADMTPT 68

QY 61 FKFEDEPKFEVLEKPOA 76  
DB 69 FKFEDEPKFEVLEKPOA 84

RESULT 3  
AAV48588  
ID AAV48588 standard; protein; 108 AA.

XX AAV48588;  
AC AAV48588;  
DT 08-DEC-1999 (first entry)

XX Human breast tumour-associated protein 49.

DE Expressed sequence tag; EST; human; breast; cancer; gene therapy;  
KM treatment; tumour; cytostatic; medicament.

XX Homo sapiens.

PN DE19813839-A1.

PD 23-SEP-1999.

XX 20-MAR-1998; 98DE-01013839.

XX 20-MAR-1998; 98DE-01013839.

PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;

DR WPI; 1999-528981/45.

DR N-PSDB; AA233657.

PT Human nucleic acid sequences and protein products from tumor breast  
PT tissue, useful for breast cancer therapy.

XX Claim 22; 164; 188pp; German.

XX This invention describes novel human nucleic acid sequences from tumor  
CC breast tissue which have cytostatic activity. The nucleic acid sequences  
CC can be used to produce and isolate full-length gene sequences. They can  
CC be used to express proteins, which can be used as tools to find an  
CC activity against breast cancer. The sequences can be used in sense or  
CC antisense form. They are especially useful for medicaments for gene  
CC therapy to treat breast cancer. AAV48540-Y48617 represent protein  
CC fragments encoded by the expressed sequence tags described in the method  
CC of the invention  
XX

SQ Sequence 108 AA;

Query Match 99.5%; Score 392; DB 2; Length 108;  
Best Local Similarity 98.7%; Pred. No. 2.9e-39;  
Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKELDPIQKLFVDKIREYKSKRQTSGGPVDASSEYQOELERELFKLKOMFGNADMTPT 60  
DB 33 NKELDPIQKLFVDKIREYKSKRQTSGGPVDASSEYQOELERELFKLKOMFGNADMTPT 92

QY 61 FKFEDEPKFEVLEKPOA 76  
DB 93 FKFEDEPKFEVLEKPOA 108

RESULT 4  
AAG04005  
ID AAG04005 standard; protein, 108 AA.  
XX  
AC AAG04005;  
XX  
DT 06-OCT-2000 (first entry)  
XX  
DE Human secreted protein, SEQ ID NO: 8086.  
XX  
KW Human, 5' EST, expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping.  
XX  
OS Homo sapiens.  
XX  
PN EPI033401-A2.  
XX  
PD 06-SEP-2000.  
XX  
PE 21-FEB-2000; 2000EP-00200610.  
XX  
PR 26-FEB-1999; 99US-0122487P.  
XX  
PA (GEST ) GENSET.  
XX  
PI Dunas Milne Edwards J, Duclert A, Giordano J;  
XX  
DR WPI; 2000-500381/45.  
DR N-PSDB; AAC04011.  
XX  
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX  
PS Claim 13; SEQ ID NO 8086; 71pp + Sequence listing; English.  
XX  
CC The present sequence is a polypeptide encoded by one of a large number of  
CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were  
CC prepared from total human RNAs or polyA+ RNAs derived from 30 different  
CC tissues. EST sequences usually correspond mainly to the 3' untranslated  
CC region (UTR) of the mRNA because they are often obtained from oligo-dT  
CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA  
CC libraries derived from the 5' ends of mRNAs and even in those cases where  
CC longer cDNA sequences have been obtained, the full 5' UTR is rarely  
CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can  
CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs  
CC are also used in diagnostic, forensic, gene therapy and chromosome  
CC mapping procedures. They are used to obtain upstream regulatory sequences  
CC and to design expression and secretion vectors  
XX  
SQ Sequence 108 AA;  
Query Match 99.5%; Score 392; DB 3; Length 108;  
Best Local Similarity 98.7%; Pred. No. 2.9e-39;  
Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NKEIDPIQKLFVDKIREYKSRQTSGGPVDASSSYQGLERELFKLKGFGNADMTPT 60  
DB 33 NKEIDPIQKLFVDKIREYKSRQTSGGPVDASSSYQGLERELFKLKGFGNADMTPT 92  
QY 61 FKFDPKFEVLEKPOA 76  
DB 93 FKFDPKFEVLEKPOA 108  
XX  
RESULT 5  
ADE63626  
ID ADE63626 standard; protein, 108 AA.  
XX  
AC ADE63626;  
XX  
DT 29-JAN-2004 (first entry)

XX  
DE Human Protein P18859, SEQ ID NO 9570.  
XX  
KW Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.  
XX  
OS Homo sapiens.  
XX  
PN WO2003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PE 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
PA (GSHO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
PI Woolf C, D'Urso D, Befort K, Costigan M;  
XX  
DR WPI; 2003-268312/26.  
DR GENBANK; P18859.  
XX  
PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
PS Claim 1; Page; 1017pp; English.  
XX  
CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 108 AA;  
Query Match 99.5%; Score 392; DB 7; Length 108;  
Best Local Similarity 98.7%; Pred. No. 2.9e-39;  
Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NKEIDPIQKLFVDKIREYKSRQTSGGPVDASSSYQGLERELFKLKGFGNADMTPT 60  
DB 33 NKEIDPIQKLFVDKIREYKSRQTSGGPVDASSSYQGLERELFKLKGFGNADMTPT 92  
QY 61 FKFDPKFEVLEKPOA 76  
DB 93 FKFDPKFEVLEKPOA 108

RESULT 6  
ADE63638  
ID ADE63638 standard; protein, 108 AA.  
XX  
XX ADE63638;  
AC  
XX 29-JAN-2004 (first entry)  
DT  
XX  
XX Human Protein P18859, SEQ ID NO 9582.  
DE  
XX  
XX Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.  
XX  
XX Homo sapiens.  
OS  
XX WO2003016475-A2.  
PN  
XX 27-FEB-2003.  
PD  
XX 14-AUG-2002; 2002WO-US025765.  
PF  
XX 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
XX (GENO) GEN HOSPITAL CORP.  
PA (FARB) BAYER AG.  
PI  
XX Woolf C, D'urso D, Befort K, Costigan M;  
XX WPI; 2003-268312/26.  
DR GENBANK; P18859.  
DR  
XX  
XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
PT  
XX  
XX Claim 1; Page; 1017pp; English.  
PS  
XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 108 AA;  
SO

Query Match 99.5%; Score 392; DB 7; Length 108;  
Best Local Similarity 98.7%; Pred. No. 2.9e-39;

Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 NKELDPIQKLFVNDKIRYKSKRQTSQGPVNASSEYOQELRELFKTKOMGNDMMTFPT 60  
Db 33 NKELDPIQKLFVNDKIRYKSKRQTSQGPVNASSEYOQELRELFKTKOMGNDMMTFPT 92  
Qy 61 FKFEDEPKFEVLEKRPQA 76  
Db 93 FKFEDEPKFEVLEKRPQA 108

RESULT 7  
ADE63634  
ID ADE63634 standard; protein, 108 AA.  
XX  
XX ADE63634;  
AC  
XX 29-JAN-2004 (first entry)  
DT  
XX  
XX Human Protein P18859, SEQ ID NO 9578.  
DE  
XX  
XX Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.  
XX  
XX Homo sapiens.  
OS  
XX WO2003016475-A2.  
PN  
XX 27-FEB-2003.  
PD  
XX 14-AUG-2002; 2002WO-US025765.  
PF  
XX 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
XX (GENO) GEN HOSPITAL CORP.  
PA (FARB) BAYER AG.  
PI  
XX Woolf C, D'urso D, Befort K, Costigan M;  
XX WPI; 2003-268312/26.  
DR GENBANK; P18859.  
DR  
XX  
XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
PT  
XX  
XX Claim 1; Page; 1017pp; English.  
PS  
XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of



CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 108 AA;

Query Match 99.5%; Score 392; DB 7; Length 108;  
Best Local Similarity 98.7%; Pred. No. 2.9e-39;  
Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NKEIDPIQKLFVDKIREYKSRQTSGGPVDASSSEYQOELERELFKLKMPGNADMTPT 60  
Db 33 NKEIDPIQKLFVDKIREYKSRQTSGGPVDASSSEYQOELERELFKLKMPGNADMTPT 92

Qy 61 FKPEDPKFEVLEKPOA 76  
Db 93 FKPEDPKFEVLEKPOA 108

RESULT 8  
ADE63630  
ID ADE63630 standard; protein; 108 AA.

AC ADE63630;  
DT 29-JAN-2004 (first entry)

DE Human Protein P18859, SEQ ID NO 9574.

XX Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KM spared nerve injury; SNI; Chung.

OS Homo sapiens.

PN WO2003016475-A2.

XX 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.

PA (GEHO ) GEN HOSPITAL CORP.

PA (FARB ) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

DR WPI; 2003-268312/26.  
DR GENBANK; P18859.

PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a

CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 108 AA;

Query Match 99.5%; Score 392; DB 7; Length 108;  
Best Local Similarity 98.7%; Pred. No. 2.9e-39;  
Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NKEIDPIQKLFVDKIREYKSRQTSGGPVDASSSEYQOELERELFKLKMPGNADMTPT 60  
Db 33 NKEIDPIQKLFVDKIREYKSRQTSGGPVDASSSEYQOELERELFKLKMPGNADMTPT 92

Qy 61 FKPEDPKFEVLEKPOA 76  
Db 93 FKPEDPKFEVLEKPOA 108

RESULT 9  
ADJ69447  
ID ADJ69447 standard; protein; 108 AA.

AC ADJ69447;

DT 06-MAY-2004 (first entry)

DE Human heat mitochondrial protein as a therapeutic target SeqID1253.

XX mitochondrial; human; screening assay; diabetes mellitus;

KW Huntington's disease; osteoarthritis;

KW Leber's hereditary optic neuropathy; LHON;

KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;

KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;

KW neuroprotective; noctropic; antidiabetic; anticonvulsant; antiepileptic;

KW osteopatinic; ophthalmological; cytoskeletal.

OS Homo sapiens.

PN WO2003087768-A2.

XX 23-OCT-2003.

PF 04-APR-2003; 2003WO-US010870.

PR 12-APR-2002; 2002US-0372843P.

PR 17-JUN-2002; 2002US-0389877P.

PR 20-SEP-2002; 2002US-0412418P.

PA (MITO-) MITOKOR.

PA (BUCK-) BUCK INST AGE RES.

PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM,

PI Warnock DE;

XX WPI; 2003-845369/78.

PT Identifying a mitochondrial target for drug screening assays and for  
PT treating diseases associated with altered mitochondrial function,  
PT comprises detecting a modified polypeptide in a sample and correlating  
PT with the disease.

XX Claim 1; SEQ ID NO 1253; 180pp; English.

XX This invention relates to novel mitochondrial targets that can be used  
CC for therapeutic intervention in treating a disease associated with  
CC altered mitochondrial function. Specifically, it refers to a method for  
CC identifying proteins of the human heart mitochondrial proteome that are  
CC useful for drug screening assays, as well as therapeutic targets. The  
CC present invention describes a method for identifying such proteins that  
CC can be used in the treatment of various diseases associated with altered  
CC mitochondrial function including diabetes mellitus, Huntington's disease,  
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial  
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy  
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these  
CC compositions have neuroprotective, neurotropic, antidiabetic,  
CC anticonvulsant, antiarrhythmic, osteopathic, ophthalmological and  
CC cytoskeletal activities. This polypeptide sequence is a human heart  
CC mitochondrial protein of the invention.

XX  
SQ Sequence 108 AA;

Query Match 99.5%; Score 392; DB 7; Length 108;  
Best Local Similarity 98.7%; Pred. No. 2.9e-39;  
Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKEIDPIQKLFVDKIREYKSKRQTSQGPVDASSEYQOELERELFKLKQMGFNADMTPT 60  
DB 33 NKEIDPIQKLFVDKIREYKSKRQTSQGPVDASSEYQOELERELFKLKQMGFNADMTPT 92  
QY 61 FKFEDEPKFEVLEKPOA 76  
DB 93 FKFEDEPKFEVLEKPOA 108

RESULT 10  
AAAG01886  
ID AAG01886 standard; protein; 106 AA.  
AC AAG01886;  
DT 06-OCT-2000 (first entry)  
XX Human secreted protein, SEQ ID NO: 5967.  
DE Human secreted protein, SEQ ID NO: 5967.  
XX Human, 5' EST, expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping.  
XX Homo sapiens.  
OS  
XX EP1033401-A2.  
PN 06-SEP-2000.  
XX 21-FEB-2000; 2000EP-00200610.  
PF 26-FEB-1999; 99US-0122487P.  
PR (GEST ) GENSET.  
XX  
PI Dumas Milne Edwards J, Duclert A, Giordano J;  
XX WPI; 2000-500381/45.  
DR N-PSDB; AAC01892.  
XX  
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX  
XX Claim 13; SEQ ID NO 5967; 71pp + Sequence Listing; English.

CC The present sequence is a polypeptide encoded by one of a large number of  
CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were  
CC prepared from total human RNAs or polyA<sup>+</sup> RNAs derived from 30 different  
CC tissues. EST sequences usually correspond mainly to the 3' untranslated  
CC region (UTR) of the mRNA because they are often obtained from oligo-dT

CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA  
CC sequences derived from the 5' ends of mRNAs and even in those cases where  
CC longer cDNA sequences have been obtained, the full 5' UTR is rarely  
CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can  
CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs  
CC are also used in diagnostic, forensic, gene therapy and chromosome  
CC mapping procedures. They are used to obtain upstream regulatory sequences  
CC and to design expression and secretion vectors

XX  
SQ Sequence 106 AA;

Query Match 97.2%; Score 383; DB 3; Length 106;  
Best Local Similarity 98.6%; Pred. No. 3.5e-38;  
Matches 73; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKEIDPIQKLFVDKIREYKSKRQTSQGPVDASSEYQOELERELFKLKQMGFNADMTPT 60  
DB 33 NKEIDPIQKLFVDKIREYKSKRQTSQGPVDASSEYQOELERELFKLKQMGFNADMTPT 92  
QY 61 FKFEDEPKFEVLEKPOA 74  
DB 93 FKFEDEPKFEVLEKPOA 106

RESULT 11  
AAAG7500  
ID AAG7500 standard; protein; 114 AA.  
AC AAG7500;  
DT 22-FEB-1999 (first entry)  
XX Human mitochondrial F6 subunit (HMF6) polypeptide.  
DE Human mitochondrial F6 subunit; HMF6; recombinant; cancer; treatment;  
KW ATP synthase coupling factor; myopathy; neurodegenerative disease.  
XX  
OS Homo sapiens.  
XX US5849527-A.  
PN 15-DEC-1998.  
PD 31-MAR-1997; 97US-00828239.  
PF 31-MAR-1997; 97US-00828239.  
PR (INCY-) INCYTE PHARM INC.  
XX  
XX Hillman JL, Shah P;  
XX WPI; 1999-069728/06.  
DR N-PSDB; AAV80300.  
XX  
XX DNA encoding human mitochondrial F6 subunit polypeptide - useful for  
PT producing recombinant polypeptide or as probe.  
XX  
XX Claim 1; Fig 1A-B; 26pp; English.

CC This represents a human mitochondrial F6 subunit (HMF6) polypeptide. Host  
CC cells comprising a vector containing the HMF6 nucleic acid can be used  
CC for the recombinant production of the protein. The HMF6 polypeptide is a  
CC novel ATP synthase coupling factor and may be used in the treatment of  
CC cancer, myopathies and neurodegenerative diseases. The nucleic acid can  
CC be used as a probe to detect HMF6 genes

XX  
SQ Sequence 114 AA;

Query Match 82.5%; Score 325; DB 2; Length 114;  
Best Local Similarity 98.4%; Pred. No. 3.9e-31;  
Matches 62; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 14 KIREYKSKRQTSQGPVDASSEYQOELERELFKLKQMGFNADMTPTFKFEDEPKFEVLEK 73

Db 52 KIREYKSKRQTSGGVVDASSEYQOELERELFKLKMFGNADNPTPTFKFEDPKFEVLEK 111  
 QY 74 POA 76  
 DB 112 POA 114

RESULT 12  
 AAY82605  
 ID AAY82605 standard; protein; 114 AA.

AC AAY82605;  
 DT 02-AUG-2000 (first entry)

DE Human mitochondrial F6 subunit protein sequence SEQ ID NO:1.

XX Human; mitochondrial F6 subunit; HMF6; cytostatic; neurotropic;  
 XX neuroprotective; antiparkinsonian; anticonvulsant; gene therapy; cancer;  
 KW myopathy; external ophthalmoplegia; Kearns-Sayre syndrome;  
 KW lactic acidosis; neurodegenerative disorder; Alzheimer's disease;  
 KW Huntington's disease; Parkinson's disease; epilepsy.

OS Homo sapiens.

PN US6048718-A.

PD 11-APR-2000.

PF 04-DEC-1998; 98US-00205679.

PR 31-MAR-1997; 97US-00828239.

PA (INCY-) INCYTE PHARM INC.

PI Shah P, Hillman JL;

DR WPI; 2000-302782/26.

DR N-PSDB; AAA13916.

XX Purified human mitochondrial F6 subunit polypeptide is useful for  
 PT treating cancer, myopathy e.g. ophthalmoplegia, Kearns-Sayre syndrome and  
 XX neurodegenerative disorders e.g. Alzheimer's disease and epilepsy.

PS Claim 1; Fig 1; 27pp; English.

XX The present sequence represents the human mitochondrial F6 subunit (HMF6)  
 CC protein. HMF6 has cytostatic, neurotropic, neuroprotective,  
 CC antiparkinsonian and anticonvulsant activities, and can be used in gene  
 CC therapy. HMF6 is useful in a composition in conjunction with a suitable  
 CC carrier for treating cancer, myopathies such as external ophthalmoplegia,  
 CC Kearns-Sayre syndrome, lactic acidosis and neurodegenerative disorders  
 CC such as Alzheimer's disease, Huntington's disease, Parkinson's disease  
 CC and epilepsy

XX Sequence 114 AA;

Query Match 82.5%; Score 325; DB 3; Length 114;  
 Best Local Similarity 98.4%; Pred. No. 3.9e-31;  
 Matches 62; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 14 KIREYKSKRQTSGGVVDASSEYQOELERELFKLKMFGNADNPTPTFKFEDPKFEVLEK 73  
 DB 52 KIREYKSKRQTSGGVVDASSEYQOELERELFKLKMFGNADNPTPTFKFEDPKFEVLEK 111

QY 74 POA 76  
 DB 112 POA 114

RESULT 13  
 AAB80964

ID AAB80964 standard; protein; 76 AA.

XX AAB80964;

AC AAB80964;  
 DT 08-JUN-2001 (first entry)

DE Rat Coupling factor 6 CF6.

XX Rat; coupling factor 6; CF6; antiinflammatory; antidiabetic; antiulcer;  
 KW cardiact; hypotensive; antianginal; antiaesthetic; antirheumatic;  
 KW antiarthritic; proton-translocating ATPase; prostaglandin;  
 KW cytoplasmic PLA<sub>2</sub>; cardiovascular disorder; diabetes;  
 KW inflammatory disorder; rheumatoid arthritis; cerebrovascular disorder.

OS Rattus sp.

PN WO200121205-A1.

PD 29-MAR-2001.

PF 03-AUG-2000; 2000WO-JP005210.

PR 17-SEP-1999; 99JP-00264687.

PA (SUNR ) SUNTORY LTD.

PI Osanai T, Magoita K;

DR WPI; 2001-266047/27.

XX Coupling factor 6 useful in diagnosis of diseases relating to excess or  
 PT lack of prostaglandin and cytoplasmic PLA<sub>2</sub> activity in blood, and in  
 XX screening drugs for treating e.g. cardiovascular infarction.

PS Claim 15; Page 56-7; 70pp; Japanese.

XX The present sequence is rat coupling factor 6 (CF6) protein. CF6 is a  
 CC subunit of a proton-translocating ATPase found in mitochondria. CF6 can  
 CC be used in diagnosis of diseases relating to excess or lack of  
 CC prostaglandin and cytoplasmic PLA<sub>2</sub> activity, and in screening inhibitors  
 CC and potentiators as drugs to treat e.g. cardiovascular infarction,  
 CC hypertension, arteriosclerosis, angina pectoris, cardiac insufficiency,  
 CC cerebrovascular disorders, hyperlipemia, diabetes, bronchitis, gastric  
 CC ulcer, pregnant eclampsia, haemolytic uremia syndrome, thrombopenic  
 CC purpura, inflammatory diseases like cerebral infarction, acute  
 CC pancreatitis, asthma, ARDS (adult respiratory distress syndrome), and  
 CC rheumatoid arthritis

XX Sequence 76 AA;

Query Match 80.5%; Score 317; DB 4; Length 76;  
 Best Local Similarity 75.0%; Pred. No. 2.2e-30;  
 Matches 57; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

QY 1 NKEIDPIOKLFVDRKIREYKSKRQTSGGVVDASSEYQOELERELFKLKMFGNADNPTPT 60  
 DB 1 NKEIDPIOKLFVDRKIREYKSKRQTSGGVVDGTPGYQOEVDRFLFKLKMFGNADNPTPT 60

QY 61 FKFEDEPKFEVLEKPOA 76  
 DB 61 FKFEDEPKFEVLEKPOA 76

RESULT 14  
 ID ADE63636  
 AC ADE63636 standard; protein; 108 AA.

AC ADE63636;

DT 29-JAN-2004 (first entry)

DE Rat Protein P21571, SEQ ID NO 9580.

KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
KM chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
XX  
OS Rattus norvegicus.  
XX  
PN WO2003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002WO-US025765.  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
PA (GENO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
PI Woolf C, D'urso D, Befort K, Costigan M,  
PI WPI; 2003-268312/26.  
DR GENBANK; P21571.  
XX  
PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
PS Claim 1; Page; 1017pp; English.  
XX

The invention discloses a composition comprising two or more isolated rat  
or human polynucleotides or a polynucleotide which represents a fragment,  
derivative or allelic variation of the nucleic acid sequence. Also  
claimed are a vector comprising the novel polynucleotide, a host cell  
comprising the vector, a method for identifying a nucleotide sequence  
which is differentially regulated in an animal subjected to pain and a  
kit to perform the method, an array, a method for identifying an agent  
that increases or decreases the expression of the polynucleotide sequence  
that is differentially expressed in neuronal tissue of a first animal  
subjected to pain, a method for identifying a compound which regulates  
the expression of a polynucleotide sequence which is differentially  
expressed in an animal subjected to pain, a method for identifying a  
compound that regulates the activity of one or more of the  
polynucleotides, a method for producing a pharmaceutical composition, a  
method for identifying a compound or small molecule that regulates the  
activity in an animal of one or more of the polypeptides given in the  
specification, a method for identifying a compound useful in treating  
pain and a pharmaceutical composition comprising the one or more  
polypeptides or their antibodies. The polynucleotide or the compound that  
modulates its activity is useful for preparing a medicament for treating  
pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
therapy). The sequence presented is a rat protein (shown in Table 2 of  
the specification) which is differentially expressed during pain. Note:  
The sequence data for this patent did not form part of the printed  
specification, but was obtained in electronic form directly from WIPO at  
ftp.wipo.int/pub/published\_pat\_sequences.

Sequence 108 AA:

Query Match 80.5%; Score 317; DB 7; Length 108;  
Best Local Similarity 75.0%; Pred. No. 3.3e-30;  
Matches 57; Conservative 10; Mismatches 9; Indels 0; Gaps 0

CY 1 NKEIDLPQKLFVKIKRIFYSKKTGSGPVADASEYGOEERELFTLKMPFGNADNNPPT 60  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |  
DB 33 NKEIDLPQOKFLVDIRIRYKANRLASGPVDITGEIVQOEVDRELFTLKNQTKGENDKFPT 92  
| | | | | FKKEDPKFEVLDEKQA 76  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |  
DB 93 FNFKEDPKFEVLDEKQKS 108

RESULT 15  
ADE63632

XX ADE63632 standard; protein; 108 AA.  
AC ADE63632;  
DT 29-JAN-2004 (first entry)  
XX  
DE Rat Protein P21571, SEQ ID NO 9576.  
XX  
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
KV chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
OS Rattus norvegicus.  
PN WO2003016475-A2.  
PP 27-FEB-2003.  
PD  
PF 14-AUG-2002; 2002MO-USO25765.  
PX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
PA (GENO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
PI Woolf C, D'Urso D, Befort K, Costigan M,  
PI WPI; 2003-268312/26.  
DR GENBANK; P21571.  
XX  
XX  
XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
PS Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (SNI) in an animal (e.g. gene therapy). The sequence presented is a rat protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Seq Sequence 108 AA;

Query Match 80.5%; Score 317; DB 7; Length 108;  
Best Local Similarity 75.0%; Pred. No. 3.3e-30;  
Matches 57; Conservative 10; Mismatches 9; Indels 0; Gaps 0

1 NKEILDPIKLPFDKREYRSKRQTSGGGVNDASSEYQGLERLELFLTKOMFGNADNNATPPT 60  
|||||:||||||:||||||:||||||:||||||:||||||:||||||:||||||:||||||:  
33 NKEILDPVQLTFRDKREYRKARLASGGGVNDGPEYQGLVDNELFLTKOMYGGENDXKPT 92

Oy	61	FKFEDPKFEVLKPOA	76
		:	
Db	93	FNFKFEDPKFEVLKPOA	108

Search completed: April 4, 2005, 19:19:22  
Job time : 127.5 secs

*This Page Blank (uspio)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2005, 18:59:30 / Search time 118.5 Seconds  
(without alignments)  
328.422 Million cell updates/sec

Title: US-09-831-951A-2  
Perfect score: 402  
Sequence: 1 NKEIDPVOKFLDKIREYKA.....KPTFNFEDPKFEVLDRKQS 76

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: UniProt\_03:\*  
1: uniprot\_sprotc:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	402	100.0	108	1	ATPR_RAT
2	358	89.1	108	1	ATPR_MOUSE
3	333	82.8	108	1	ATPR_BOVIN
4	331	82.3	108	1	Q8SPH6
5	329	81.8	76	1	ATPR_PIG
6	315	78.4	108	1	ATPR_HUMAN
7	306	76.1	108	2	Q6N259
8	241	60.0	107	2	Q6PG55
9	238	59.2	107	2	Q6ERJ1
10	198	49.3	112	2	Q6NYF7
11	162	40.3	117	2	Q7Q7P4
12	156.5	38.9	106	1	ATPR_DROME
13	132.5	33.0	147	2	Q9VZ72
14	106	26.4	54	2	Q812D0
15	80.5	20.0	285	2	Q6R840
16	77.5	19.3	578	2	Q05924
17	76.5	19.0	219	2	Q750B1
18	75.5	18.8	1654	2	Q8FM04
19	75	18.7	1454	2	Q7WZK4
20	74	18.4	143	2	Q8YVU3
21	74	18.4	1196	2	Q8PMW0
22	72.5	18.0	858	2	Q17647
23	72.5	18.0	860	2	Q95NM4
24	71.5	17.8	1405	2	Q22691
25	71	17.7	382	2	Q93516
26	70.5	17.5	262	2	Q82HA2
27	70.5	17.5	1934	2	Q75UR1
28	70	17.4	570	2	Q7UH95
29	69	17.4	1103	2	Q64XG8
30	68.5	17.0	1242	2	Q6CNK0
31	68	16.9	210	2	Q6C3F2

32	68	16.9	385	2	Q8K687	Q8K687 streptococ
33	68	16.9	1919	2	Q7RRU1	Q7RRU1 plasmodium
34	68	16.9	2719	2	Q7RP00	Q7RP00 plasmodium
35	67	16.7	322	2	Q6APB7	Q6APB7 desulfotale
36	67	16.7	797	2	Q67651	Q67651 aquilex aeo
37	67	16.7	914	2	Q8S5Q7	Q8S5Q7 oryza sativ
38	67	16.7	976	2	Q6FKX5	Q6FKX5 candida gla
39	67	16.7	996	2	Q24303	Q24303 plasmu sctiv
40	67	16.7	1140	2	Q7PDQ7	Q7PDQ7 plasmodium
41	66.5	16.5	425	2	Q6MMQ3	Q6MMQ3 bdellovibri
42	66.5	16.5	458	2	Q81US8	Q81US8 homo sapien
43	66.5	16.5	531	2	Q6UWV6	Q6UWV6 homo sapien
44	66.5	16.5	531	2	Q6ZTS5	Q6ZTS5 homo sapien
45	66	16.4	101	2	Q8EN19	Q8EN19 oceanobacil

## ALIGNMENTS

```

RESULT 1
ATPR_RAT      STANDARD;      PRT;      108 AA.
ID   ATPR_RAT
AC   P21571;
DT   01-MAY-1991 (Rel. 18, Created)
DT   01-MAY-1991 (Rel. 18, Last sequence update)
DT   05-JUN-2004 (Rel. 44, Last annotation update)
DE   ATP synthase coupling factor 6, mitochondrial precursor (EC 3.6.3.14)
DE   (Fe6).
GN   Name=Atp5f1;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=91024964; PubMed=2145831;
RA   Higuti T., Osaka F., Yoshihara Y., Tsurumi C., Kawamura Y., Tani I.,
RA   Toda H., Kakuno T., Sakiyama F., Tanaka K., Ichihara A.;
RT   "cDNA cloning and sequencing for the import precursor of coupling
RT   factor 6 in H(+)-ATP synthase from rat liver mitochondria.";
RL   Biochem. Biophys. Res. Commun. 171:1079-1086(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Anterior pituitary;
RX   MEDLINE=92339904; PubMed=1386054; DOI=10.1016/0378-1119(92)90528-W;
RA   Tracer H.L., Ioh Y.P., Birch N.P.;
RT   "Rat mitochondrial coupling factor 6: molecular cloning of a cDNA
RT   encoding the imported precursor.";
RL   Gene 116:291-292(1992).
RN   [3]
RP   SEQUENCE OF 33-66.
RC   TISSUE=Liver;
RX   MEDLINE=93054567; PubMed=1429613;
RA   Higuti T., Yoshihara Y., Kuroiwa K., Kawamura Y., Toda H.,
RA   Sakiyama F.;
RT   "A simple, rapid method for purification of epsilon-subunit, coupling
RT   factor 6, subunit d, and subunit e from rat liver H(+)-ATP synthase
RT   and determination of the complete amino acid sequence of epsilon-
RT   subunit.";
RL   J. Biol. Chem. 267:22658-22661(1992).
RN   [4]
RP   FUNCTION: This is one of the chains of the nonenzymatic component
RP   (CF0) subunit of the mitochondrial ATPase complex. F6 seems to
RP   be part of the stalk that links CF(0) to CF(1).
RN   [5]
RP   CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
RP   H(+) (Out).
RN   [6]
RP   SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
RP   core - and CF(0) - the membrane proton channel. CF(0) seems to
RP   have nine subunits: a, b, c, d, e, f, g, F6 and 8 (or A6L).
RN   [7]
RP   This SWISS-PROT entry is copyright. It is produced through a collaboration
RP   between the Swiss Institute of Bioinformatics and the EMBL outstation -
RP   the European Bioinformatics Institute. There are no restrictions on its
  
```

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M73030; AAA40954.1; -;  
 DR EMBL; X54510; CA38369.1; -;  
 DR PIR; J01167; J01167.  
 DR HSC-2DPAGE; P21571; RAT.  
 DR RGD; 621376; Alp5j.  
 DR InterPro: IPR00387; ATP synth\_F6.  
 DR Pfam; PF05511; ATP-synt\_F6; 1.  
 KW CF(0); Direct protein sequencing; Hydrogen ion transport;  
 FT Mitochondrion; Transit peptide.  
 FT TRANSLIT 1 32  
 FT CHAIN 33 108 ATP synthase coupling factor 6.  
 SQ SEQUENCE 108 AA; 12494 MW; FFI177C9681B5F51 CRC64;

Query Match 100.0%; Score 402; DB 1; Length 108;  
 Best Local Similarity 100.0%; Pred. No. 1,6e-34;  
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKEIDPVOKFLDKIRREYKAKRLASGGPVDTGPEYQOEVDRELFLKQMGKGMKXPT 60  
 DB 33 NKEIDPVOKFLDKIRREYKAKRLASGGPVDTGPEYQOEVDRELFLKQMGKGMKXPT 92

QY 61 FNFEDPKFEVLDKPOS 76  
 DB 93 FNFEDPKFEVLDKPOS 108

RESULT 2  
 ATPR MOUSE STANDARD; PRT; 108 AA.  
 ID ATPR MOUSE  
 AC P97450;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE ATP synthase coupling factor 6, mitochondrial precursor (EC 3.6.3.14)  
 DE (Fe).  
 GN Name=Atpsj;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6;  
 RA Rocha D., Anderson E., Botcherby M., Jordan B., Carrier A.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6j;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nakai I., Osato N., Saito R., Suzuki H., Yamana I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,  
 RA Baldarelli R., Hill D.P., Bull C., Hume D.A., Quackenbush J.,  
 RA Schiraldi L.M., Kanapin A., Matsumura H., Batalov S., Bensei K.W.,  
 RA Blake J.A., Bratt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S.,  
 RA Gaasterland T., Gariboldi M., Giesli C., Godzik A., Gough J.,  
 RA Grimmond S., Guclincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konegaya A., Kurochkin I.V., Lee Y., Lemhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Nimata K., Okido T., Pavan W.J., Pereira G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.V., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verrato R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilting L.G., Wynnshaw-Boris A., Yanagisawa M., Yang I., Yang L.,

RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kikukawa T., Kono H., Nakamura M., Sakazume N., Sato K.,  
 RA Ritschke T., Waki K., Kawai J., Aizawa K., Arkawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs";  
 RL Nature 420:563-573(2002).

RN [3].  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Colon;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Millay S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek U.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyik S.W.,  
 RA Villalón D.R., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Wilting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,  
 RA Butlerfield Y.S.N., Krzywicki M.I., Skalska U., Smalins D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC - FUNCTION: This is one of the chains of the nonenzymatic component  
 CC (CF(0) subunit) of the mitochondrial ATPase complex. F6 seems to  
 CC be part of the stalk that links CF(0) to CF(1).  
 CC - CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +  
 CC H(+) (Out).  
 CC - SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic  
 CC core - and CF(0) - the membrane proton channel. CF(0) seems to  
 CC have nine subunits: a, b, c, d, e, f, g, F6 and 8 (or A6L).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U77128; AAB19213.1; -;  
 DR EMBL; AK078484; BAC37301.1; -;  
 DR EMBL; BC010766; AAH10766.1; -;  
 DR SWISS-2DPAGE; P97450; MOUSE.  
 MG; MG1:107777; Alp5j.  
 DR InterPro: IPR006187; ATP synth\_F6.  
 DR Pfam; PF05511; ATP-synt\_F6; 1.  
 KW CF(0); Hydrogen ion transport; Mitochondrion; Transit peptide.  
 FT TRANSLIT 1 32  
 FT CHAIN 33 108 Mitochondrion (By similarity).  
 FT ATP synthase coupling factor 6.  
 SQ SEQUENCE 108 AA; 12496 MW; E2A2E63F723C8BF CRC64;

Query Match 89.1%; Score 356; DB 1; Length 108;  
 Best Local Similarity 85.5%; Pred. No. 6,8e-30;  
 Matches 65; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 1 NKEIDPVOKFLDKIRREYKAKRLASGGPVDTGPEYQOEVDRELFLKQMGKGMKXPT 60  
 DB 33 NKEIDPVOKFLDKIRREYKAKRLASGGPVDTGPEYQOEVDRELFLKQMGKGMKXPT 92

QY 61 FNFEDPKFEVLDKPOS 76



Db 93 FKPDPPKREVIDKPOS 108

## RESULT 3

ATPR\_BOVIN STANDARD; PRT; 108 AA.  
 AC P02721;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE ATP synthase coupling factor 6, mitochondrial precursor (EC 3.6.3.14) (Fe).  
 DE Name=ATP5F;  
 GN Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.  
 OC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88163536; PubMed=2894843;  
 RA Walker J.E., Gay N.J., Powell S.O., Kostina M., Dyer M.R.;  
 RT "ATP synthase from bovine mitochondria: sequences of imported  
 RT precursors of oligomycin sensitivity conferral protein, factor 6, and  
 RT adenosine triphosphatase inhibitor protein.";  
 RL Biochemistry 26:8613-8619(1987).  
 RN [2]  
 RP SEQUENCE OF 33-108.  
 RX MEDLINE=85038563; PubMed=6149548;  
 RA Fang J.-K., Jacobs J.W., Kanner B.I., Racker R., Bradshaw R.A.;  
 RT "Amino acid sequence of bovine heart coupling factor 6.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 81:6603-6607(1984).  
 RN [3]  
 RP SEQUENCE OF 33-37.  
 RC TISSUE=Heart;  
 RX MEDLINE=91242449; PubMed=1827992;  
 RA Walker J.E., Lutter R., Dupuis A., Runswick M.J.;  
 RT "Identification of the subunits of F1F0-ATPase from bovine heart  
 RT mitochondria.";  
 RL Biochemistry 30:5369-5378(1991).  
 CC -1- FUNCTION: This is one of the chains of the nonenzymatic component  
 CC (CF(0) subunit) of the mitochondrial ATPase complex. F6 seems to  
 CC be part of the stalk that links CF(0) to CF(1). Also involved in  
 CC the restoration of oligomycin-sensitive ATPase activity to  
 CC depleted F1F0 complexes.  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +  
 CC H(+) (Out).  
 CC -1- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic  
 CC core - and CF(0) - the membrane proton channel. CF(0) seems to  
 CC have nine subunits: a, b, c, d, e, f, g, F6 and 8 (or 8eL).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by, and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M19217; AAA30511.1; -.  
 DR PIR; B27382; JIB06.  
 DR InterPro; IPR008387; ATP\_synth\_F6.  
 DR Pfam; PF05511; ATP-synth\_F6\_1.  
 KW CF(0); Direct protein sequencing; Hydrogen ion transport;  
 KM Mitochondrion; Transit peptide.  
 FT TRANSIT 1 32  
 FT CHAIN 33 108 ATP synthase coupling factor 6.  
 FT CONFLICT 94 94 T->F (in Ref. 2).  
 SQ SEQUENCE 108 AA; 12532 MW; E5376A0518C3E1C8 CRC64;

Query Match 82.8%; Score 333; DB 1; Length 108;  
 Best Local Similarity 78.9%; Pred. No. 2.9e-27;  
 Matches 60; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

Qy 1 NKELDPVQKLFVDKIREYKAKRLASGGPVDTPGPEYQOEVDRELFLKQMYGKGEMDKRPT 60  
 Db 33 NKELDPVQKLFVDKIREYKAKRLASGGPVDTPGPEYQOEVDRELFLKQMYGKGEMDKRPT 92  
 Qy 61 FNFEDPKREVIDKPOS 76  
 Db 93 FTFEDPKREVEKPOS 108

## RESULT 4

Q8SPH6 PRELIMINARY; PRT; 108 AA.  
 ID O8SPH6  
 AC O8SPH6;  
 DT 01-JUN-2002 (TRENBLREL. 21, Created)  
 DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)  
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)  
 DE ATP synthase subunit F6.  
 GN Name=ATP5F;  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 OC NCBI\_TaxID=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain parietal lobe;  
 RA Osada N., Kusuda J., Hirata M., Tanuma R., Hida M., Sugano S.,  
 RA Hirai M., Hashimoto K.;  
 RT "Search for genes positively selected during primate evolution by 5'-  
 RT end-sequence screening of cynomolgus monkey cDNAs.";  
 RL Genomics 79:657-662(2002).  
 RN EMBL; AB072025; BAB6814.1; -.  
 DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.  
 DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.  
 DR GO; GO:0046961; F:hydrogen-transporting ATP synthase activity; IEA.  
 DR GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.  
 DR InterPro; IPR008387; ATP\_synth\_F6.  
 DR Pfam; PF05511; ATP-synth\_F6\_1.  
 SQ SEQUENCE 108 AA; 12587 MW; A33AC44891716C1E CRC64;

Query Match 82.3%; Score 331; DB 2; Length 108;  
 Best Local Similarity 79.7%; Pred. No. 4.7e-27;  
 Matches 59; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

Qy 1 NKELDPVQKLFVDKIREYKAKRLASGGPVDTPGPEYQOEVDRELFLKQMYGKGEMDKRPT 60  
 Db 33 NKELDPVQKLFVDKIREYKAKRLASGGPVDTPGPEYQOEVDRELFLKQMYGKGEMDKRPT 92  
 Qy 61 FNFEDPKREVIDKPOS 74  
 Db 93 FTFEDPKREVEKPOS 106

## RESULT 5

ATPR\_PIG STANDARD; PRT; 76 AA.  
 ID ATPR\_PIG  
 AC P13618;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE ATP synthase coupling factor 6, mitochondrial (EC 3.6.3.14) (Fe).  
 GN Name=ATP5F;  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OC NCBI\_TaxID=99823;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Intestine;  
 RX MEDLINE=88083634; PubMed=2961617; DOI=10.1016/0014-5793(87)80547-1;  
 RA Chen Z.-W., Mutt V., Barros-Scoederling J., Joernvall H.;  
 RT "Isolation and structural characterization of porcine coupling factor

```

RT 6 from intestinal tissues."
RL FEBS Lett. 226:43-46(1987).
CC -1- FUNCTION: This is one of the chains of the nonenzymatic component
CC (CF(0) subunit) of the mitochondrial ATPase complex. F6 seems to
CC be part of the stalk that links CF(0) to CF(1).
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
CC H(+) (out).
CC -1- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
CC core - and CF(0) - the membrane proton channel. CF(0) seems to
CC have nine subunits: a, b, c, d, e, f, g, F6 and 8 (or A6L).
DR PIR: S00212; S00212.
DR Interpro: IPR008387; ATP_synth_F6.
DR Pfam: PF05511; ATP-synth_F6; 1.
DR CF(0); Direct protein sequencing; Hydrogen ion transport;
KM Mitochondrion.
SQ SEQUENCE 76 AA; 8930 MW; 16F73EE42F87B874 CRC64;

Query Match 81.8%; Score 329; DB 1; Length 76;
Best Local Similarity 77.6%; Pred. No. 5.2e-27;
Matches 59; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Dy 1 NKEIDPQKFLDKIRRYKAKRLASGPPVDTGPYQGEVRELFKLGKMGKEMKEPT 60
Dy 1 NKEIDPQKFLVDKIRRYKRTKQTSQGPVDAQPEYQDLDRELFKLGKMGKADMTFPN 60
Dy 61 NFEDPKFEVLDKPKQS 76
Dy 61 FTEDPKFEAVEKPKQS 76
Dy 61 FTEDPKFEAVEKPKQS 76

RESULT 6
ATPR_HUMAN STANDARD; PRT; 108 AA.
ID ATPR_HUMAN
AC P18859;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE ATP synthase coupling factor 6, mitochondrial precursor (EC 3.6.3.14)
DE (F6).
GN Name=ATP6; Synonyms=ATPM;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RX MEDLINE=9115366; PubMed=1825642; DOI=10.1016/0378-1119(91)90068-W;
RA Uaved A.A., Ogata K., Sanadi D.R.;
RT "Human mitochondrial ATP synthase: cloning cDNA for the nuclear-
RT encoded precursor of coupling factor 6.";
RL Gene 97:307-310(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=9115366; PubMed=1830479;
RA Hagiuti T., Tsurumi C., Kawamura Y., Tsujita H., Osaka F.,
RA Yoshikawa Y., Tani I., Tanaka K., Ichihara A.;
RT "Molecular cloning of cDNA for the import precursor of human coupling
RT factor 6 of H(+) -ATP synthase in mitochondria.";
RL Biochem. Biophys. Res. Commun. 178:793-799(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Kidney;
RX MEDLINE=91154917; PubMed=11230166; DOI=10.1101/gr.154701;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glaesl S.,
RA Kistner W., Boecker M., Bloeker H., Baerends S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koerner K., Strack N.,
RA Mewes H.-W., Ottensmeyer F., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs.";
RL Genome Res. 11:422-435(2001).
RN [4]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=20289799; PubMed=10690953; DOI=10.1038/35012518;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Seoda E.,
RA Onki H., Takagi T., Sakaki Y., Tauden S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh T., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Kordaslek G., Hornischer K., Brandt P.,
RA Scharte M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Risseimann L., Degen E.,
RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Leberich H., Reinhardt R., Vasp M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
RN [5]
RP SEQUENCE FROM N.A.
RX TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uadin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Millamy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]
RP SEQUENCE OF 33-43.
RX TISSUE=Liver;
RX MEDLINE=93162045; PubMed=1286669;
RA Hochstrasser D.F., Frutiger S., Paquet N., Bairoch A., Ravier F.,
RA Pasquali C., Sanchez J.-C., Tissot J.-D., Bjellqvist B., Vargias R.,
RA Appel R.D., Hughes G.O.;
RT "Human liver protein map: a reference database established by
RT microsequencing and gel comparison.";
RL Electrophoresis 13:992-1001(1992).
CC -1- FUNCTION: This is one of the chains of the nonenzymatic component
CC (CF(0) subunit) of the mitochondrial ATPase complex. F6 seems to
CC be part of the stalk that links CF(0) to CF(1). Also involved in
CC the restoration of oligomycin-sensitive ATPase activity to
CC depleted F1-F0 complexes.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
CC H(+) (out).
CC -1- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
CC core - and CF(0) - the membrane proton channel. CF(0) seems to
CC have nine subunits: a, b, c, d, e, f, g, F6 and 8 (or A6L).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M37104; AAA51807.1; -
DR EMBL: W73031; AAA58630.1; -
DR EMBL: AL110183; CAB53667.1; -
DR EMBL: AP001694; -; NOT ANNOTATED_CDS.
DR EMBL: BC001178; AAH01178.1; -

```

DR PIR; JT0563; JT0563.  
DR SWISS-2DPAGE; P18859; HUMAN.  
DR Genew: HGNC:847; ATP5J.  
DR H-InVDB: H1X0016040; -.  
DR Reactome: P18859; -.  
DR MIM: 603152; -.  
DR GO; GO:0005743; C:mitochondrial inner membrane; TAS.  
DR GO; GO:0005739; C:mitochondrion; TAS.  
DR InterPro: IPR008387; ATP synth F6.  
DR Pfam: PF05511; ATP-synt\_F6; 1\_F6.  
DR CF(O); Direct protein sequencing; Hydrogen ion transport;  
KM Mitochondrion; Transit peptide;  
FT TRANSIT 1 32 Mitochondrion.  
FT CHAIN 33 108 ATP synthase coupling factor 6.  
FT CONFLICT 68 68 O -> H (in Ref. 2).  
SQ SEQUENCE 108 AA; 12587 MW; EDC1A4F01A10F17 CRC64;  
  
Query Match 78.4%; Score 315; DB 1; Length 108;  
Best Local Similarity 73.7%; Pred. No. 2,3e-25;  
Matches 56; Conservative 11; Mismatches 9; Indels 0; Gaps 0;  
  
QY 1 NKELDPIQKFLDIREYKAKRLASGGPVDTGPEYQGVNDRELFKLKMVGKGMKDKPPT 60  
DB 33 NKELDPIQKFLVDIREYKSKRQTSGGPVDSBYQQLERLEFKLKMFGNADMTFTPT 92  
QY 61 FNFEDPKFEVLDRKPOS 76  
DB 93 FKPEDPKFEVLEKPKQA 108  
  
RESULT 7  
Q6NZ59 PRELIMINARY; PRT; 108 AA.  
AC Q6NZ59;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE ATP synthase, H+ transporting, mitochondrial F0 complex, subunit F6,  
DE isoform a.  
GN Name=ATP5J;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OC NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,  
RA Krzywnski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Klein S., Strausberg R.,  
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC066310; AAH6310.1; -.  
DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.  
DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.  
DR GO; GO:0046961; F:hydrogen-transporting ATPase activity; IEA.  
DR GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.  
DR InterPro: IPR008387; ATP synth F6.  
DR Pfam: PF05511; ATP-synt\_F6; 1\_F6.  
SQ SEQUENCE 108 AA; 12627 MW; EDD2CDDCF01A10F17 CRC64;  
  
Query Match 76.1%; Score 306; DB 2; Length 108;  
Best Local Similarity 72.4%; Pred. No. 2e-24;  
Matches 55; Conservative 11; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 NKELDPIQKFLDIREYKAKRLASGGPVDTGPEYQGVNDRELFKLKMVGKGMKDKPPT 60  
DB 33 NKELDPIQKFLVDIREYKSKRQTSGGPVDSBYQQLERLEFKLKMFGNADMTFTPT 92  
QY 61 FNFEDPKFEVLDRKPOS 76  
DB 93 FKPEDPKFEVLEKPKQA 108  
  
RESULT 8  
Q6PG55 PRELIMINARY; PRT; 107 AA.  
AC Q6PG55;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE MGC68738 protein.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
OC Xenopodinae; Xenopus.  
OC NCBI\_Taxid=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,  
RA Krzywnski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.;  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative";  
RL Dev. Dyn. 225:384-391(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;  
RA Klein S., Strausberg R.,  
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL: BC057213; AAH57213.1; -.  
 DR GO: GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.  
 DR GO: GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.  
 DR GO: GO:0046961; F:hydrogen-transporting ATPase activity; IEA.  
 DR GO: GO:0015986; P:ATP synthesis coupled proton transport; IEA.  
 DR InterPro: IPR008387; ATP synth\_F6.  
 DR Pfam: PF05511; ATP-synt F6; 1.  
 DR SQUENCE 107 AA; 12359 MW; 0F2C5BD80E6F286 CRC64;

Query Match 60.0%; Score 241; DB 2; Length 107;  
 Best Local Similarity 61.8%; Pred. No. 1.3e-17;  
 Matches 42; Conservative 13; Mismatches 13; Indels 0; Gaps 0;

QY 2 KELDVPQKFLDKIRRYKAKRLASGSPVDTGPYQOEVDRELFLKQMYGKGMDFKPF 61  
 DB 36 KELDVPQKFLDKIRRYKAKRLASGSPVDTGPYQOEVDRELFLKQMYGKGMDFKPF 95  
 QY 62 NFEDPKFE 69  
 DB 96 KFEPPKFE 103

RESULT 9  
 Q68FJ1 PRELIMINARY; PRT; 107 AA.  
 AC 068FJ1  
 DT 25-OCT-2004 (TREMBlrel. 28, Created)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DE MG86324 protein.  
 GN Name=MG86324;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xeropodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22341133; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S.L., Strusberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.,  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 initiative."  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX PubMed=12477932; DOI=10.1073/pnae.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.P., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Klein S., Gerhard D.S.;

RL Submitted (Aug-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC079787; AAH79787.1; -.  
 DR InterPro: IPR008387; ATP synth\_F6.  
 DR Pfam: PF05511; ATP-synt F6; 1.  
 DR SQUENCE 107 AA; 12364 MW; C6EBBC7EF08B54B CRC64;

Query Match 59.2%; Score 238; DB 2; Length 107;  
 Best Local Similarity 61.8%; Pred. No. 2.8e-17;  
 Matches 42; Conservative 13; Mismatches 13; Indels 0; Gaps 0;

QY 2 KELDVPQKFLDKIRRYKAKRLASGSPVDTGPYQOEVDRELFLKQMYGKGMDFKPF 61  
 DB 36 KELDVPQKFLDKIRRYKAKRLASGSPVDTGPYQOEVDRELFLKQMYGKGMDFKPF 95  
 QY 62 NFEDPKFE 69  
 DB 96 KFEPPKFE 103

RESULT 10  
 Q6N7F7 PRELIMINARY; PRT; 112 AA.  
 AC 06N7F7  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Zgc:77541.  
 GN ORFNames=zgc:77541;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnae.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.P., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX PubMed=12477932; DOI=10.1073/pnae.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.P., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Klein S., Gerhard D.S.;

Query Match 49.3%; Score 198; DB 2; Length 112;  
 QY 2 KELDVPQKFLDKIRRYKAKRLASGSPVDTGPYQOEVDRELFLKQMYGKGMDFKPF 61  
 DB 36 KELDVPQKFLDKIRRYKAKRLASGSPVDTGPYQOEVDRELFLKQMYGKGMDFKPF 95  
 QY 62 NFEDPKFE 69  
 DB 96 KFEPPKFE 103

Best Local Similarity 56.5%; Pred. No. 4,7e-13;  
Matches 35; Conservative 13; Mismatches 14; Indels 0; Gaps 0;  
QY 2 KELDPOVKLFLDKIREYKAKRLASGSPVDPGPEYQOQVDERELFLKQMYGKMGENDKRPPT 61  
Db 37 KMDPIQKLFUDKIRIDYNSKSKASGVADGVPYQKNLAETTYLQRYGGGDLSPQPF 96  
QY 62 NF 63  
Db 97 SF 98  
RESULT 11  
ID 0707P4 PRELIMINARY; PRT; 117 AA.  
AC 0707P4;  
DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
DE ACP445 (Fragment).  
GN Name=agCG56939; ORFNames=ENSGNG0000019097;  
OS Anopheles gambiae str. PEST.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.  
OX NCBI\_TaxID=180454;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PEST;  
RL Anopheles Genome Sequencing Consortium;  
RA Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.  
CC -1- CAUTION: The sequence shown here is derived from an  
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; AAB0108952; EAA10598.1;  
DR GO; GO:0016469; C:Proton-transporting two-sector ATPase complex; IEA.  
DR GO; GO:0046933; F:Hydrogen-transporting ATP synthase activity; IEA.  
DR GO; GO:0046961; F:Hydrogen-transporting ATPase activity; rota. .; IEA.  
DR GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.  
DR InterPro: IPR008387; ATP\_synth\_F6.  
DR Pfam: PF05511; ATP\_synth\_F6; 1.  
FT NON TER 1  
SQ SEQUENCE 117 AA; 12984 MW; 98DASB5F9376B95 CRC64;  
Query Match 40.3%; Score 162; DB 2; Length 117;  
Best Local Similarity 44.0%; Pred. No. 3e-09;  
Matches 33; Conservative 16; Mismatches 22; Indels 4; Gaps 2;  
QY 2 KELDPOVKLFLDKIREYKAKRLASGSPVDPGPEYQOQVDERELFLKQMYGKMGENDKRPPT 57  
Db 40 KASPIQQLFVTKIRVYAKQSGAGKLVDAIPEIQRELEKQEMKAKQYGGGGEPMYTA 99  
QY 58 PPTNFEDPKFEVLD 72  
Db 100 FPAKFEEPKIDPIN 114  
RESULT 12  
ID ATPR\_DROME STANDARD; PRT; 106 AA.  
AC 024407; O9VCN0;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 25-JAN-2005 (Rel. 46, Last annotation update)  
DE ATP synthase coupling factor 6, mitochondrial precursor (EC 3.6.3.14)  
DE (FE).  
GN Name=ATP6vtn-Cf6; ORFNames=CG4412;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.

RC TISSUE=Ovary;  
RX MEDLINE=99168769; PubMed=10071211; DOI=10.1007/s004380050942;  
RA Casaghe C., Ragone G., Perrini B., Moschetti R., De Pinto V.,  
RA Catizzi R., Barsanti P.;  
RT "Identification of nuclear genes encoding mitochondrial proteins:  
RT isolation of a collection of D. melanogaster cDNAs homologous to  
RT sequences in the Human Gene index database.";  
RL Mol. Genet. 261:64-70(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkely;  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Durbin K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris K.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostali M., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lei Z.,  
RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye Y., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RL "The genome sequence of Drosophila melanogaster.";  
Science 287:2185-2195(2000).  
RN [3]  
RP GENOME REANNOTATION.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Miara S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
RA Smith C.D., Tupy U.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Belencourt B.R., Celniker S.E., de Grey A.D.N.J., Dysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
RT systematic review.";  
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
CC -1- FUNCTION: This is one of the chains of the nonenzymatic component  
CC (CF(0) subunit) of the mitochondrial ATPase complex. F6 seems to  
CC be part of the stalk that links CF(0) to CF(1).  
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +  
CC H(+) (Out).  
CC -1- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic  
CC core - and CF(0) - the membrane proton channel. CF(0) seems to  
CC have nine subunits: a, b, c, d, e, f, g, F6 and 8 (or A6L).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

CC -----  
 DR EMBL; X99665; CA67979.1; -  
 DR EMBL; AE003743; AAF56127.1; -  
 DR IntAct; Q24407; -  
 DR FlyBase; FBgn0016119; ATPsyn-Ctf6.  
 DR InterPro; IPR008387; ATP\_synth\_F6.  
 DR Pfam; PF05511; ATP\_synth\_F6; 1.  
 CC CE(0); Hydrogen ion transport; Mitochondrion; Transmembrane peptide.  
 FT TRANSIT 1  
 FT CHAIN ? 106 ATP synthase coupling factor 6.  
 FT SEQUENCE 106 AA; 11936 MW; 08E1B074EB34E9AB CRC64;

Query Match 38.9%; Score 156.5; DB 1; Length 106;  
 Best Local Similarity 44.3%; Pred. No. 1e-08;  
 Matches 35; Conservative 15; Mismatches 24; Indels 5; Gaps 3;

QY 1 NKEIDPQKFLDKIRRYAKRLASGPGVDTGPEYQOEVDRELFLKQKMT---GKGMMDK 57  
 DB 29 NKASDPIQQLFLDKIRRYAKRLASGPGVDTGPEYQOEVDRELFLKQKMT---GKGMMDK 87

QY 58 FPFNFEDPKFB-VLDKRPQ 75  
 DB 88 FPFNFEDPKFB-VLDKRPQ 106

RESULT 13  
 ID Q9VZ72 PRELIMINARY; PRT; 147 AA.  
 AC Q9VZ72;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE CG12027-PA.  
 GN ORFNames=CG12027;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NC NCB1\_Taxid=7227;  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazer V.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L.,  
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Berens P.V., Bernier B.P., Bhandari D., Bolshakov S.,  
 RA Borkov A., Borchan M.R., Bouck J., Brockstein P., Brottier P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo R., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson R.K., Dou P.L., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fester C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibsen C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris A., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodgett, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao S., Zhu X., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426055; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
 RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila  
 RT melanogaster euchromatic genome sequence.";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the Drosophila melanogaster euchromatin:  
 RT a genomic perspective.";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426099; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tuzy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt R.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX FlyBase;  
 RG Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE003483; AAF4954.1; -  
 DR FlyBase; FBgn0035585; CG12027.  
 DR GO; GO:0016699; C:proton-transporting two-sector ATPase complex; IEA.  
 DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.  
 DR GO; GO:0046961; F:hydrogen-transporting ATPase activity; IEA.  
 DR GO; GO:0015986; P:ATP synthase coupled proton transport; IEA.  
 DR InterPro; IPR008387; ATP\_synth\_F6.  
 DR Pfam; PF05511; ATP\_synth\_F6; 1.  
 DR SEQUENCE 147 AA; 16749 MW; B63D1117EB785CD CRC64;

Query Match 33.0%; Score 132.5; DB 2; Length 147;  
 Best Local Similarity 39.7%; Pred. No. 4.9e-06;  
 Matches 31; Conservative 12; Mismatches 28; Indels 7; Gaps 3;

QY 5 DPVQKFLDKIRRYAKRLASGPGVDTGPEYQOEVDRELFLKQKMT---MKPFPF 61  
 DB 26 DPVQKFLDKIRRYAKRLASGPGVDTGPEYQOEVDRELFLKQKMT---MKPFPF 84

OY 62 NFED--PKFVLDKPKQS 76  
Db 85 KLPDIDIDIPISVDLPEN 102

RESULT 14

Q812D0 PRELIMINARY; PRT; 54 AA.  
AC Q812D0;  
DT 01-JUN-2003 (TREMBlrel. 24, Created)  
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE ATP synthase coupling factor VI (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxId=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=15199140; DOI=10.1128/MCB.24.13.5844-5849.2004;  
RA Risteovski S., O'Leary D.A., Thornell A.P., Owen M.J., Koia I.,  
RA Hertzog P.J.;  
RT "The ETS Transcription Factor GABP(alpha) is Essential for Early  
RT Embryogenesis.";  
RL Mol. Cell. Biol. 24:5844-5849(2004).  
DR EMBL; AF346288; AAC27831.1; -  
DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.  
DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.  
DR GO; GO:0046961; F:hydrogen-transporting ATP synthase activity; IEA.  
DR GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.  
DR InterPro; IPR008387; ATP\_synth\_F6.  
DR Pfam; PF05511; ATP-synt\_F6; 1.  
FT NON TER 54  
SQ SEQUENCE 54 AA; 6285 MW; 82BFF0707AA7A29C CRC64;

Query Match 26.4%; Score 106; DB 2; Length 54;  
Best Local Similarity 90.9%; Pred. No. 0.00099;  
Matches 20; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKEIDPVOKLFLDKIREYKAKR 22  
Db 33 NKEIDPVOKLFLDKIREYKSKR 54

RESULT 15

Q6R840 PRELIMINARY; PRT; 285 AA.  
AC Q6R840;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE 770RFP09.  
OS Bacteriophage 77.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.  
OX NCBI\_TaxId=259901;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=14716317;  
RA Liu J., Dehbi M., Moock G., Arhin F., Bauda P., Bergeron D.,  
RA Callejo M., Ferretti V., Ha N., Kwan T., McCarty J., Srikanth R.,  
RA Williams D., Wu J.J., Gros P., Pelletier J., Dubow M.;  
RT "Antimicrobial drug discovery through bacteriophage genomics.";  
DR EMBL; AY508466; AAR67881.1; -  
SQ SEQUENCE 285 AA; 32577 MW; 67798DFDE282AE16 CRC64;

Query Match 20.0%; Score 80.5; DB 2; Length 285;  
Best Local Similarity 30.6%; Pred. No. 3;  
Matches 26; Conservative 14; Mismatches 20; Indels 25; Gaps 4;

OY 4 LDPVOKLFLDKIREYKAKRLASGSPVDTGPE-----YQGEVDRELKLT 46  
Db 1 MDPILGKIDIKIIEG-----ASKGPVETFSKTWELVFGKFLYVDKVIYQREVEFE--KF 53

OY 47 KQMYGKGEKDPTEFNEFDPKREVL 71  
Db 54 KEQF-KKEISSVPKNNLQEPQSL 77

Search completed: April 4, 2005, 19:23:26  
Job time : 120.5 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2005, 19:11:13 ; Search time 26 Seconds  
(without alignments)  
281.249 Million cell updates/sec

Title: US-09-831-951A-2

Perfect score: 402  
Sequence: 1 NKEUDPVKLFIDKIREYKA.....KPTFNFEDPKFEVLDPKQOS 76

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	402	100.0	108	2 UC1167	coupling factor 6
2	333	82.8	108	1 UJB06	coupling factor 6
3	329	81.8	76	2 S00212	coupling factor 6,
4	315	78.4	108	2 J70563	coupling factor 6
5	77.5	19.3	578	2 S51379	probable phosphoes
6	74	18.4	143	2 A12057	hypothetical prote
7	73	18.2	15	2 EP0444	coupling factor 6
8	72.5	18.0	858	2 T18946	probable phospholi
9	71.5	17.8	813	2 T25104	hypothetical prote
10	71	17.7	253	2 T12008	hypothetical prote
11	67	16.7	797	2 A70453	glutamate ammonia
12	66	16.4	497	2 A84641	hypothetical prote
13	65	16.2	257	2 G70453	indole-3-glycerol
14	65	16.2	443	1 D64584	heat shock protein
15	65	16.2	443	2 F71929	heat shock protein
16	64.5	16.0	285	2 J00448	CATG-binding facto
17	64.5	16.0	1940	1 A24922	myosin heavy chain
18	64.5	16.0	1940	1 S04090	myosin heavy chain
19	64	15.9	808	2 C72858	Acort-66 protein -
20	63.5	15.8	244	2 T25970	hypothetical prote
21	63.5	15.8	486	2 T24334	hypothetical prote
22	63.5	15.8	557	2 A61256	myosin heavy chain
23	63.5	15.8	996	2 S71750	import intermediat
24	63.5	15.8	1934	2 I48153	myosin heavy chain
25	63.5	15.8	1935	1 A57102	myosin beta heavy
26	63.5	15.8	1938	1 S06005	myosin alpha heavy
27	63.5	15.8	1938	1 I49464	alpha cardiac myos
28	63.5	15.8	1939	1 A46762	myosin alpha heavy
29	63.5	15.8	1939	2 I48175	myosin heavy chain

30	63	15.7	364	2 T00200	hypothetical prote
31	63	15.7	1613	2 G64488	reverse gyrase (in
32	62.5	15.5	129	2 T11904	hypothetical prote
33	62.5	15.5	302	2 S66751	single stranded D
34	62.5	15.5	353	1 S66750	single stranded D
35	62.5	15.5	761	2 E64449	hypothetical prote
36	62.5	15.5	1935	1 S06006	myosin beta heavy
37	62	15.4	259	2 D85088	protein R09F10.1 (
38	62	15.4	500	2 A43385	glutaryl-tRNA (gln)
39	62	15.4	1177	2 I64233	hypothetical prote
40	61.5	15.3	231	2 T32070	hypothetical prote
41	61.5	15.3	233	2 S70531	btk2.11 protein pr
42	61.5	15.3	390	2 S56560	hypothetical 43.6k
43	61.5	15.3	390	2 B91291	hypothetical prote
44	61.5	15.3	390	2 D86132	hypothetical prote
45	61	15.2	450	2 A11259	signal recognition

## ALIGNMENTS

## RESULT 1

UC1167  
coupling factor 6 precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 09-Oct-1992 #sequence revision 09-Oct-1992 #text\_change 09-Jul-2004  
C:Accession: UC1167, A35933, C44300, I52239, S23283  
R:Tracer, H.L.; Loh, Y.P.; Birch, N.P.  
Gene 16, 291-292, 1992  
A:Title: Rat mitochondrial coupling factor 6: molecular cloning of a cDNA encoding the  
A:Reference number: JC1167, MUID:92339904, PMID:136054  
A:Accession: UC1167  
A:Molecule type: mRNA  
A:Residues: 1-108 <TR>  
A:Cross-references: UNIPROT:P21571; EMBL:X54510; NID:g14214; PID:CA38369.1; PID:G5995  
R:Hiquti, T.; Osaka, F.; Yoshinara, Y.; Tsutsumi, C.; Kawamura, Y.; Tan, I.; Toda, H.;  
Biochem. Biophys. Res. Commun. 171, 1079-1086, 1990  
A:Title: cDNA cloning and sequencing for the import precursor of coupling factor 6 in f  
A:Reference number: A55933; MUID:91024964; PMID:2145831  
A:Accession: A55933  
A:Molecule type: mRNA  
A:Residues: 1-108 <TR>  
A:Cross-references: GB:M73030; NID:G203529; PID:AAA40954.1; PID:G203530  
A>Note: part of this sequence, including the amino end of the mature protein, was deter  
R:Hiquti, T.; Yoshinara, Y.; Kuroiwa, K.; Kawamura, Y.; Toda, H.; Sakiyama, F.  
J. Biol. Chem. 267, 22658-22661, 1992  
A:Title: A simple, rapid method for purification of epsilon-subunit, coupling factor 6,  
of epsilon-subunit.  
A:Reference number: A44300; MUID:93054567; PMID:1429613  
A:Accession: C44300  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 33-66 <HI2>  
A:Experimental source: liver  
A>Note: sequence extracted from NCBI backbone (NCBIP:119665)  
C:Comment: This is one of the soluble components required for coupling of phosphorylati  
C:Superfamily: coupling factor 6  
C:Keywords: mitochondrion; oxidative phosphorylation  
P:1-32/Domain: transit peptide (mitochondrion) #status predicted <TMP>  
P:33-108/Product: coupling factor 6 #status experimental <MAT>

Query Match	100.0%	Score 402,	DB 2,	Length 108;
Best Local Similarity	100.0%	Pred. No. 1.9e-36;		
Matches	76;	Conservative 0;	Mismatches 0;	Gaps 0;
Oy	1	NKEUDPVKLFIDKIREYKARLASGGPVDTPGPEYQOEVDRELFKLQMYGKGMDEKPEPT	60	
Db	33	NKEUDPVKLFIDKIREYKARLASGGPVDTPGPEYQOEVDRELFKLQMYGKGMDEKPEPT	92	
Oy	61	FNFPDPKFEVLDPKQOS 76		
Db	93	FNFPDPKFEVLDPKQOS 108		

## RESULT 2

J1B06

coupling factor 6 precursor, mitochondrial - bovine

N/Alternate names: ATP synthase coupling factor 6

C/Species: Bos primigenius taurus (cattle)

C/Date: 17-May-1995 #sequence\_revision 28-Oct-1994 #text\_change 09-Jul-2004

C/Accession: B27382; A03180; E39566

R/Walker, J.E.; Gay, N.J.; Powell, S.J.; Kostina, M.; Dyer, M.R.

A/Title: ATP synthase from bovine mitochondria: sequences of imported precursors of O1g

A/Reference number: A90527; MUID:88163536; PMID:2894843

A/Accession: B27382

A/Molecule type: mRNA

A/Residues: 1-108 &lt;WML&gt;

A/Cross-references: UNIPROT:P02721; GB:M19217; NID:g163035; PIDN:AAA30511.1; PID:g163036

R/Fang, J.; Jacob, J.W.; Kanner, B.I.; Recker, E.; Bradshaw, R.A.

Proc. Natl. Acad. Sci. U.S.A. 81, 6603-6607, 1984

A/Title: Amino acid sequence of bovine heart coupling factor 6.

A/Reference number: A03180; MUID:85038563; PMID:6149548

A/Accession: A03180

A/Molecule type: protein

A/Residues: 33-93; P, 95-108 &lt;FAN&gt;

A/Experimental source: heart

R/Malker, J.E.; Lutter, R.; Dupuis, A.; Runswick, M.J.

Biochemistry 30, 5369-5378, 1991

A/Title: Identification of the subunits of F<sub>1</sub>-F<sub>0</sub>-ATPase from bovine heart mitochondria.

A/Reference number: A39566; MUID:91244459; PMID:1827992

A/Accession: E39566

A/Molecule type: protein

A/Residues: 33-37 &lt;MA2&gt;

C/Comment: This is one of the soluble components required for coupling of phosphorylation

C/Superfamily: coupling factor 6

C/Keywords: mitochondrial; oxidative phosphorylation

F<sub>1</sub>-F<sub>32</sub>/Domain: transist peptide (mitochondrion) #status predicted <TMP>F<sub>33</sub>-108/Product: coupling factor 6 #status experimental <MAT>

Query Match

Best Local Similarity 82.8%; Score 333; DB 1; Length 108;

Matches 60; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 1 NKEIDPVQKFLDKIRRYKAKRLASGGPVDTPGPEYQGVDRLEFKLKQYKGMKMDKPT 60

DB 33 NKEIDPVQKFLDKIRRYKAKRLASGGPVDTPGPEYQGVDRLEFKLKQYKGMKMDKPT 92

QY 61 FNEEDPKFEVLDRKPOS 76

DB 93 FTEDPKFEVEKPOS 108

## RESULT 3

S00212

coupling factor 6, mitochondrial - pig

C/Species: Sus scrofa domestica (domestic pig)

C/Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 09-Jul-2004

C/Accession: S00212

R/Chen, Z.W.; Mutt, V.; Barros-Soederling, J.; Joernvall, H.

FEBS Lett. 226, 43-46, 1987

A/Title: Isolation and structural characterization of porcine coupling factor 6 from int

A/Reference number: S00212; MUID:88083634; PMID:2961617

A/Accession: S00212

A/Molecule type: protein

A/Residues: 1-76 &lt;CHE&gt;

A/Residues: 1-76 &lt;CHE&gt;

A/Cross-references: UNIPROT:P13618

C/Superfamily: coupling factor 6

C/Keywords: mitochondrial; oxidative phosphorylation

Query Match

Best Local Similarity 81.8%; Score 329; DB 2; Length 76;

Matches 59; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 NKEIDPVQKFLDKIRRYKAKRLASGGPVDTPGPEYQGVDRLEFKLKQYKGMKMDKPT 60

DB 33 NKEIDPVQKFLDKIRRYKAKRLASGGPVDTPGPEYQGVDRLEFKLKQYKGMKMDKPT 92

QY 61 FNEEDPKFEVLDRKPOS 76

DB 93 FTEDPKFEVEKPOS 108

DB 1 NKEIDPVQKFLDKIRRYKAKRLASGGPVDTPGPEYQGVDRLEFKLKQYKGMKMDKPT 60

QY 61 FNEEDPKFEVLDRKPOS 76

DB 61 FTEDPKFEVEKPOS 76

## RESULT 4

J10563

coupling factor 6 precursor, mitochondrial - human

N/Alternate names: ATP synthase coupling factor 6

C/Species: Homo sapiens (man)

C/Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 09-Jul-2004

C/Accession: J10563; J01066; T14747

R/Atwood, R.A.; Ogata, K.; Sanadi, D.R.

Gene 97, 307-310, 1991

A/Title: Human mitochondrial ATP synthase: cloning cDNA for the nuclear-encoded precurs

A/Reference number: J10563; MUID:91153664; PMID:1825642

A/Accession: J10563

A/Molecule type: mRNA

A/Residues: 1-108 &lt;JAV&gt;

A/Cross-references: UNIPROT:P18859; GB:M37104; NID:g179274; PIDN:AAA51807.1; PID:g179274

A/Experimental source: fetal muscle

R/Higuti, T.; Tsurumi, C.; Kawamura, Y.; Teujita, H.; Osaka, F.; Yoshitara, Y.; Tani, I

Biochem. Biophys. Res. Commun. 178, 793-799, 1991

A/Title: Molecular cloning of cDNA for the import precursor of human coupling factor 6

A/Reference number: J01066; MUID:91315516; PMID:1830479

A/Accession: J01066

A/Molecule type: mRNA

A/Residues: 1-67; H', 69-108 &lt;HIG&gt;

A/Cross-references: GB:M73031; NID:g183785; PIDN:AAA58630.1; PID:g183786

A/Experimental source: kidney

R/Bium, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, August 1999

A/Reference number: Z18179

A/Accession: T14747

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-108 &lt;BLU&gt;

A/Cross-references: EMBL:AL110183

A/Experimental source: fetal kidney; clone DKFZ566A221

C/Genetics:

A/Gene: GDB:ATP5J

A/Cross-references: GDB:127519

A/Map position: 2pter-2qter

A/Note: DKFZ566A221.1

C/Superfamily: coupling factor 6

C/Keywords: mitochondrial; oxidative phosphorylation

F<sub>1</sub>-F<sub>32</sub>/Domain: transist peptide (mitochondrion) #status predicted <TPP>F<sub>33</sub>-108/Product: coupling factor 6 #status predicted <MAT>

Query Match

Best Local Similarity 78.4%; Score 315; DB 2; Length 108;

Matches 56; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

QY 1 NKEIDPVQKFLDKIRRYKAKRLASGGPVDTPGPEYQGVDRLEFKLKQYKGMKMDKPT 60

DB 33 NKEIDPVQKFLDKIRRYKAKRLASGGPVDTPGPEYQGVDRLEFKLKQYKGMKMDKPT 92

QY 61 FNEEDPKFEVLDRKPOS 76

DB 93 FTEDPKFEVEKPOS 108

## RESULT 5

S51379

probable phosphoesterase (EC 3.1.-.-) YJR361c [similarity] - Yeast (Saccharomyces cerev

N/Alternate names: hypothetical protein I6039.13

C/Species: Saccharomyces cerevisiae

C/Date: 23-Feb-1995 #sequence\_revision 12-May-1995 #text\_change 16-Aug-2004

C/Accession: S51379

R/Du, Z.

submitted to the EMBL Data Library, December 1994

QY 1 NKEIDPVQKFLDKIRRYKAKRLASGGPVDTPGPEYQGVDRLEFKLKQYKGMKMDKPT 60

DB 33 NKEIDPVQKFLDKIRRYKAKRLASGGPVDTPGPEYQGVDRLEFKLKQYKGMKMDKPT 92

QY 61 FNEEDPKFEVLDRKPOS 76

DB 93 FTEDPKFEVEKPOS 108

A:Description: The sequence of *S. cerevisiae* cosmid 8039.  
A:Reference number: S51377  
A:Accession: S51379  
A:Molecule type: DNA  
A:Residues: 1-578 <DUZ>  
A:Cross-references: UNIPROT:O05824; EMBL:U19103; NID:g609404; PID:g609417; GSPDB:GN00012  
C:Comment: This sequence has motifs characteristic of a variety of phosphoesterases.  
C:Genetics:  
A:Gene: MIPS:YLR361c  
A:Cross-references: SGD:S0004353  
A:Map position: 12R  
C:Superfamily: phosphoesterase core homology  
C:Keywords: hydrolase  
F:249-339/Domain: phosphoesterase core homology <PEC>

Query Match 19.3%; Score 77.5; DB 2; Length 578;  
Best Local Similarity 33.3%; Pred. No. 1.7;  
Matches 30; Conservative 10; Mismatches 29; Indels 21; Gaps 6;

QY 7 VQKLF-LDKIR-----EYKAKRLASGPDVTGPEYQOVDRELFLKQW----YKGE--- 54  
Db 206 LQKFPPLVKYVSEBAYLTTFVGRPYDTGAQRQETDEGKFKIVQLADLHLGVSESEC 265  
QY 55 MDKPPTNF--EDPKFEV-----LDKPQ 75  
Db 266 IDEYPKHEACKADPKETFTFVQVLDIERQ 295

RESULT 6  
A12057  
hypothetical protein alr2015 [imported] - *Noctoc* sp. (strain PCC 7120)  
C:Species: *Noctoc* sp. PCC 7120  
A:Note: *Noctoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C:Accession: A12057  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasaemoto, S.; Watanabe, A.; Iriyuchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena*  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: A12057  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-143 <KUR>  
A:Cross-references: UNIPROT:Q8VYG3; GB:BA000019; PIDN:BAW73714.1; PID:g17131105; GSPDB:C  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: alr2015

Query Match 18.4%; Score 74; DB 2; Length 143;  
Best Local Similarity 30.8%; Pred. No. 0.78;  
Matches 28; Conservative 9; Mismatches 18; Indels 36; Gaps 6;

QY 5 DPV-OKFLDKIRRYKAKRLASGPGVDTG-----EYQOVDRELFLKQWYKGM 55  
Db 28 DPVQSTF-----YR-----ALGGVDRGERSIALAKRFQBEIQAEINIRYL---GCM 74  
QY 56 DKPPTNF-----FEDPKFEVD 72  
Db 75 ENLFTFNARKGHEIIQLYQCDPAPKPHOLE 105

RESULT 7  
PD0444  
coupling factor 6 mitochondrial - mouse (fragment)  
C:Species: *Mus musculus* (house mouse)  
C:Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 05-Feb-1999  
C:Accession: PD0444  
R:Kawakami, T.; Uchida, T.; Sakai, T.; Kamo, M.; Morimasa, T.; Tsugita, A.  
Submitted to JIPID, August 1998  
A:Description: Proteome analysis of mouse brain.  
A:Reference number: PD0441  
A:Contents: Striatum

A:Accession: PD0444  
A:Molecule type: protein  
A:Residues: 1-15 <KAW>  
C:Keywords: mitochondrion

Query Match 18.2%; Score 73; DB 2; Length 15;  
Best Local Similarity 93.3%; Pred. No. 0.072;  
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKEIDPVOKFLDKI 15  
Db 1 NKEIDPVOKFLDKI 15

RESULT 8  
T18946  
probable phospholipase activating protein C05C10.6 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T18946; T24252  
R:Matthews, P.  
Submitted to the EMBL Data Library, February 1995  
A:Reference number: Z19049  
A:Accession: T18946  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-858 <WIL>  
A:Cross-references: UNIPROT:Q17647; EMBL:Z48178; PIDN:CAA88206.1; GSPDB:GN00020; CESP:C  
A:Experimental source: clone C05C10  
R:Wilkinson, J.  
Submitted to the EMBL Data Library, October 1995  
A:Reference number: Z19863  
A:Accession: T24252  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-858 <W12>  
A:Cross-references: EMBL:Z66515; PIDN:CAA91354.1; GSPDB:GN00020; CESP:C05C10.6  
A:Experimental source: clone R53  
C:Genetics:  
A:Gene: CESP:C05C10.6  
A:Map position: 2  
A:Introns: 15/3; 120/1; 155/3; 407/3; 513/1; 549/1; 593/3; 711/2; 786/3; 821/3

Query Match 18.0%; Score 72.5; DB 2; Length 858;  
Best Local Similarity 35.8%; Pred. No. 9.2;  
Matches 24; Conservative 8; Mismatches 26; Indels 9; Gaps 3;

QY 4 LDPVOKFLDKIRRYKAKRLASGPGVDT-GPEYQOVDRELFLKQWYKGMDFPTFN 62  
Db 399 LEETTRFKDRIPEARAFDLKSGKKVIVDGKEYD-----YALGVNFGKGPDPKQMPFN 451  
QY 63 F-EDPKF 68  
Db 452 VNESPQF 458

RESULT 9  
T25144  
hypothetical protein T22H6.6 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T25144  
R:McMurray, A.  
Submitted to the EMBL Data Library, August 1995  
A:Reference number: Z19986  
A:Accession: T25144  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-613 <WTL>  
A:Cross-references: UNIPROT:Q22691; EMBL:Z50797; PIDN:CAA90676.1; GSPDB:GN00028; CESP:T  
A:Experimental source: clone T22H6  
C:Genetics:  
A:Gene: CESP:T22H6.6

A:Map position: X  
A:introns: 93/1; 132/1; 170/1; 193/3; 209/1; 246/3; 345/3; 445/3; 490/3; 522/3; 593/1; 6

Query Match 17.8%; Score 71.5; DB 2; Length 813;  
Best Local Similarity 26.4%; Pred. No. 11;  
Matches 19; Conservative 18; Mismatches 30; Indels 5; Gaps 3;

Qy 1 NKEIDPVQKFLDKIREYKAKRLASGSPVDTGPEYQOEVDRELFLKQMGKGMDEKPT 60  
Db 744 HKEVFPRTNALILKIRVQRMRMSGSAE--PSYKPTIENITIALIDYK--PHENKLP 799  
Qy 61 FNPEDPKEVD 72  
Db 800 SSFSS-STELID 810

RESULT 10  
T21008  
hypothetical protein F16B12.8 - *Caenorhabditis elegans*  
C/Species: *Caenorhabditis elegans*  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T21008  
R:Alnecough, R.  
submitted to the EMBL Data Library, October 1996  
A/Reference number: Z19359  
A/Accession: T21008  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-253 <WII>  
A/Cross-references: UNIPROT:Q93516; EMBL:Z81064; PIDN:CAB02960.1; GSPDB:GNO0028; CESP:FL  
A/Experimental source: clone F16B12  
C/Genetics:  
A:Gene: CESP:F16B12.8  
A:Map position: X  
A:introns: 26/3; 63/3; 148/1

Query Match 17.7%; Score 71; DB 2; Length 253;  
Best Local Similarity 36.5%; Pred. No. 3.2;  
Matches 19; Conservative 7; Mismatches 26; Indels 0; Gaps 0;

Qy 1 NKEIDPVQKFLDKIREYKAKRLASGSPVDTGPEYQOEVDRELFLKQMGK 52  
Db 202 NMDLQYVMFMQFMDPKFKRLSDTFVAGAYSNRNDRDVKKYMFGK 253

RESULT 11  
A70453  
glutamate ammonia ligase adenyl-yl-transferase - *Aquifex aeolicus*  
C/Species: *Aquifex aeolicus*  
C/Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
C/Accession: A70453  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O  
V.  
Nature 392, 353-358, 1998  
A/Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.  
A/Reference number: A70300; MUID:98196666; PMID:9537320  
A/Accession: A70453  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-797 <AQP>  
A/Cross-references: UNIPROT:O67651; GB:AE000755; NID:92984063; PIDN:AAC07604.1; PID:9298  
A/Experimental source: strain VFS  
C/Genetics:  
A:Gene: glne

Query Match 16.7%; Score 67; DB 2; Length 797;  
Best Local Similarity 35.5%; Pred. No. 33;  
Matches 27; Conservative 8; Mismatches 25; Indels 16; Gaps 5;

Qy 1 NKEIDPV-----QKFLDKIREYKAKRLASGSPVDTGPEYQOEVDRELFLKQMGK 53  
Db 553 NSDIDLVFAVKSIEK--EKVHE-KAKELVRLTAHTTEGYLSVD--FRLPMGSKG 605

Qy 54 EMDKPTFNEDPKFE 69  
Db 606 EL--VFIFDYKXKFE 619

RESULT 12  
A84641  
hypothetical protein At2g24830 (imported) - *Arabidopsis thaliana*  
C/Species: *Arabidopsis thaliana* (mouse-ear cress)  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C/Accession: A84641  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentol, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L  
euse, D.; Nierman, W.C.; White, O.; Eichen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,  
Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A/Reference number: A84420; MUID:20083487; PMID:10617197  
A/Accession: A84641  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-497 <STO>  
A/Cross-references: UNIPROT:Q9SK49; GB:AE002093; NID:94559368; PIDN:AAD23029.1; GSPDB:GT  
A/Genes: At2g24830  
A:Gene: At2g24830  
A:Map position: 2

Query Match 16.4%; Score 66; DB 2; Length 497;  
Best Local Similarity 31.1%; Pred. No. 24;  
Matches 19; Conservative 11; Mismatches 23; Indels 8; Gaps 3;

Qy 1 NKEIDPVQKFLDKIREYKAKRLASGSPVDTGP--EYQOEVD--DBELFLKQMGKGM 55  
Db 396 NEGIFPRH--EKHSESVKRNQKGPVDRKALVEYQOEVDRLKLEMLKLEQVNRNKK 452

Qy 56 D 56  
Db 453 D 453

RESULT 13  
G70453  
indole-3-glycerol phosphate synthase - *Aquifex aeolicus*  
C/Species: *Aquifex aeolicus*  
C/Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
C/Accession: G70453  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O  
V.  
Nature 392, 353-358, 1998  
A/Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.  
A/Reference number: A70300; MUID:98196666; PMID:9537320  
A/Accession: G70453  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-257 <AQP>  
A/Cross-references: UNIPROT:O67657; GB:AE000756; NID:92984076; PIDN:AAC07616.1; PID:929  
A/Experimental source: strain VFS  
C/Genetics:  
A:Gene: trpC  
C/Superfamily: indole-3-glycerol-phosphate synthase; trpC homology  
F/3-254/Domain: trpC homology <TRC>

Query Match 16.2%; Score 65; DB 2; Length 257;  
Best Local Similarity 30.0%; Pred. No. 15;  
Matches 12; Conservative 12; Mismatches 10; Indels 6; Gaps 1;

Qy 11 FLDKIREYKAKRLASGSPVDTGPEYQOEVDRELFLKQMG 50  
Db 3 FLEVRSYKESQ-----IDTSPEYLRKLELIBERKERY 36

RESULT 14  
D64584  
heat shock protein - *Helicobacter pylori* (strain 2695)

C.Species: *Helicobacter pylori*  
C.Date: 10-Sep-1999 #sequence 10-Sep-1999 #text\_change 09-Jul-2004  
C.Accession: D64584  
R.Tomb, J.F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khairak, H.G., Glodek, A., McEnnenney, U.D., Kelley, J.M., Cotton, M.D., Weidman, J.M., Fujii, C., Bowman, C., Wathey, L. Nature 368, 539-547, 1997  
A.Author: Wallin, E., Hayes, W.S., Borodovsky, M., Karpk, P.D., Smith, H.O., Fraser, C. A.TITLE: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.  
A.Reference number: A64520; MUID:97394467; PMID:925185  
A.Accession: D64584  
A.Status: preliminary; nucleic acid sequence not shown; translation not shown  
A.Molecule type: DNA  
A.Residues: 1-443 <TOM>  
A.Cross-references: UNIPROT:Q5254; GB:AE000566; GB:AE000511; NID:G2313628; PIDN:AD0755  
C.Superfamily: heat shock protein hsp70; FRSB/SECL8/CDC48-type ATP-binding domain homolog  
C.Keywords: nucleotide binding; P-loop  
F:59-66/Region: nucleotide-binding motif A (P-loop)

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2005, 18:46:29 ; Search time 30 Seconds  
(without alignments)  
189.111 Million cell updates/sec

Title: US-09-831-951A-2

Perfect score: 402  
Sequence: 1 NKEHDPVQKFLDKIREYKA.....KFTFNPEPKFEVLDRKQS 76

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTCUTS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	402	100.0	108	2	US-08-828-239-5
2	402	100.0	108	3	US-09-205-679-5
3	333	82.8	108	2	US-08-828-239-4
4	333	82.8	108	3	US-09-205-679-4
5	315	78.4	108	4	US-09-513-999C-8086
6	315	78.4	108	4	US-09-949-016-11049
7	310	77.1	108	2	US-08-828-239-3
8	310	77.1	108	3	US-09-205-679-3
9	309	76.9	106	4	US-09-513-999C-5967
10	252	62.7	114	2	US-08-828-239-1
11	252	62.7	114	3	US-09-205-679-1
12	156	38.8	69	4	US-09-513-999C-5968
13	67	16.7	583	4	US-09-489-039A-7653
14	64.5	16.0	1940	4	US-09-538-092-901
15	64.5	16.0	1963	4	US-09-949-016-8888
16	63.5	15.8	460	4	US-09-438-917-17
17	63.5	15.8	460	4	US-09-949-016-6596
18	63.5	15.8	1118	4	US-09-949-016-10404
19	63.5	15.8	1886	4	US-08-938-105-3
20	63.5	15.8	1935	4	US-09-538-092-916
21	63.5	15.8	1939	3	US-09-310-187A-1
22	63.5	15.8	1939	4	US-09-538-092-917
23	63.5	15.8	1944	4	US-09-949-016-10929
24	63	15.7	351	4	US-09-489-039A-7442
25	62.5	15.5	606	4	US-09-252-991A-17409
26	62	15.4	539	4	US-09-248-796A-16770
27	61.5	15.3	290	4	US-09-107-433-3900

28	61.5	15.3	420	4	US-09-489-039A-8961	Sequence 8961, Ap
29	61	15.2	98	4	US-09-270-767-36028	Sequence 36028, A
30	61	15.2	98	4	US-09-270-767-51245	Sequence 51245, A
31	61	15.2	765	4	US-09-538-092-1340	Sequence 1340, Ap
32	60.5	15.0	555	4	US-09-543-681A-7166	Sequence 7166, Ap
33	60.5	15.0	1937	4	US-09-538-092-918	Sequence 918, App
34	60	14.9	392	4	US-09-198-452A-933	Sequence 933, App
35	60	14.9	507	4	US-09-949-016-11504	Sequence 11504, A
36	60	14.9	688	4	US-09-438-185A-870	Sequence 870, App
37	59.5	14.8	106	4	US-09-107-532A-5357	Sequence 5357, App
38	59.5	14.8	167	4	US-09-270-767-62466	Sequence 62466, A
39	59.5	14.8	454	4	US-09-270-767-46837	Sequence 46837, A
40	59.5	14.8	473	4	US-09-328-352-4658	Sequence 4658, Ap
41	59.5	14.8	690	4	US-09-489-039A-11040	Sequence 11040, A
42	59.5	14.8	970	2	US-08-449-645A-11	Sequence 11, App1
43	58.5	14.8	970	2	US-08-702-367A-11	Sequence 11, App1
44	59.5	14.8	970	5	PCT-US95-04681-11	Sequence 11, App1
45	59.5	14.8	994	3	US-08-542-635-2	Sequence 2, App11

#### ALIGNMENTS

RESULT 1  
US-08-828-239-5  
; Sequence 5, Application US/08828239  
; Patent No. 5849527  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: NOVEL ATP SYNTHASE COUPLING FACTOR 6  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/828,239  
; FILING DATE: Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PR-0260 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 108 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: Genbank  
; CLONE: 203530  
; US-08-828-239-5  
Query Match 100.0%; Score 402; DB 2; Length 108;  
Best Local Similarity 100.0%; Pred. No. 1,6e+43;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKEIDPVOKLFLDKIREYKAKRLASGGPVDGTGPEYQOEVDRELFKLKQMYGKGMDEKFP 60  
Db 33 NKEIDPVOKLFLDKIREYKAKRLASGGPVDGTGPEYQOEVDRELFKLKQMYGKGMDEKFP 92  
QY 61 FNFEDPKFEVLDKPOS 76  
Db 93 FNFEDPKFEVLDKPOS 108

## RESULT 2

US-09-205-679-5  
Sequence 5, Application US/09205679  
Patent No. 6048718  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
TITLE OF INVENTION: NOVEL ATP SYNTHASE COUPLING FACTOR 6  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/205,679  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/828,239  
FILING DATE: 03/31/1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0260 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 108 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: Genbank  
CLONE: 203530  
US-09-205-679-5

Query Match 100.0%; Score 402; DB 3; length 108;

Best Local Similarity 100.0%; Pred. No. 1.6e-43;

Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKEIDPVOKLFLDKIREYKAKRLASGGPVDGTGPEYQOEVDRELFKLKQMYGKGMDEKFP 60  
Db 33 NKEIDPVOKLFLDKIREYKAKRLASGGPVDGTGPEYQOEVDRELFKLKQMYGKGMDEKFP 92  
QY 61 FNFEDPKFEVLDKPOS 76  
Db 93 FNFEDPKFEVLDKPOS 108

## RESULT 3

US-08-828-239-4  
Sequence 4, Application US/08828239

Patent No. 5849527  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
TITLE OF INVENTION: NOVEL ATP SYNTHASE COUPLING FACTOR 6  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/828,239  
FILING DATE: Herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0260 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 108 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: Genbank  
CLONE: 625323  
US-08-828-239-4

Query Match 82.8%; Score 333; DB 2; length 108;

Best Local Similarity 78.9%; Pred. No. 9.2e-35;

Matches 60; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 1 NKEIDPVOKLFLDKIREYKAKRLASGGPVDGTGPEYQOEVDRELFKLKQMYGKGMDEKFP 60  
Db 33 NKEIDPVOKLFLDKIREYKAKRLASGGPVDGTGPEYQOEVDRELFKLKQMYGKGMDEKFP 92  
QY 61 FNFEDPKFEVLDKPOS 76  
Db 93 FNFEDPKFEVLDKPOS 108

## RESULT 4

US-09-205-679-4  
Sequence 4, Application US/09205679  
Patent No. 6048718  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
TITLE OF INVENTION: NOVEL ATP SYNTHASE COUPLING FACTOR 6  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304



COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/205,679  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/828,239  
FILING DATE: 03/31/1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0260 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 108 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 625323  
US-09-205-679-4

Query March 82.8%; Score 333; DB 3; Length 108;  
Best Local Similarity 78.9%; Pred. No. 9.2e-35;  
Matches 60; Conservative 8; Mismatches 8; Indels 0; Gaps 0;  
QY 1 NKEIDPVQKFLDKIREYKAKRLASGGPVDTPGPEYQOEVDRELFLKQMYGKGMNDKFPPT 60  
DB 33 NKEIDPVQKFLVDKIREYKAKRLASGGPVDAGPEYQOELRELFLKQMYGKGMNDKFPPT 92  
QY 61 FNEEDPKFEVLDKPKQS 76  
DB 93 FTFEDPKFEVLEKPKQS 108

RESULT 5  
US-09-513-999C-8086  
Sequence 8086, Application US/09513999C  
Patent No. 6783961  
GENERAL INFORMATION:  
APPLICANT: Dumesnil, J.B.  
APPLICANT: Dumesnil, J.B.  
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
Patent No. 6783961  
FILE REFERENCE: 59, US2, REG  
CURRENT APPLICATION NUMBER: US/09/513,999C  
CURRENT FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/122,487  
PRIOR FILING DATE: 1999-02-26  
NUMBER OF SEQ ID NOS: 36681  
SOFTWARE: Patent.pm  
SEQ ID NO 8086  
LENGTH: 108  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-513-999C-8086

Query March 78.4%; Score 315; DB 4; Length 108;  
Best Local Similarity 73.7%; Pred. No. 1.8e-32;  
Matches 56; Conservative 11; Mismatches 9; Indels 0; Gaps 0;  
QY 1 NKEIDPVQKFLDKIREYKAKRLASGGPVDTPGPEYQOEVDRELFLKQMYGKGMNDKFPPT 60  
DB 33 NKEIDPVQKFLVDKIREYKAKRLASGGPVDAGPEYQOELRELFLKQMYGKGMNDKFPPT 92

DB 33 NKEIDPVQKFLVDKIREYKAKRLASGGPVDASSEYQOELRELFLKQMYGKGMNDKFPPT 92  
QY 61 FNEEDPKFEVLDKPKQS 76  
DB 93 FTFEDPKFEVLEKPKQA 108

RESULT 6  
US-09-949-016-11049  
Sequence 11049, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: C0001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11049  
LENGTH: 108  
TYPE: PRT  
ORGANISM: Human  
US-09-949-016-11049

Query March 78.4%; Score 315; DB 4; Length 108;  
Best Local Similarity 73.7%; Pred. No. 1.8e-32;  
Matches 56; Conservative 11; Mismatches 9; Indels 0; Gaps 0;  
QY 1 NKEIDPVQKFLDKIREYKAKRLASGGPVDTPGPEYQOEVDRELFLKQMYGKGMNDKFPPT 60  
DB 33 NKEIDPVQKFLVDKIREYKAKRLASGGPVDASSEYQOELRELFLKQMYGKGMNDKFPPT 92  
QY 61 FNEEDPKFEVLDKPKQS 76  
DB 93 FTFEDPKFEVLEKPKQA 108

RESULT 7  
US-08-828-239-3  
Sequence 3, Application US/08828239  
Patent No. 5849527  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Hillman, Jennifer L.  
TITLE OF INVENTION: NOVEL ATP SYNTHASE COUPLING FACTOR 6  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/828,239  
FILING DATE: Herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:

STRANDEDNESS: single

ZIP: 94304  
COMPUTER READABLE FORM:

STRANDEDNESS: single

ZIP: 94304  
COMPUTER READABLE FORM:

```

;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PasterSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/828,239
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0260 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: 2496341
;
US-08-828-239-1

Query Match      62.7%; Score 252; DB 2; Length 114;
Best Local Similarity 71.4%; Pred. No. 1.9e-24;
Matches 45; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY      14 KIRRYKAKRLASGGPVDTGPTGYQGVDRBELFKLQMTYKSGMDKFPPTNFEDPKFEVIDK 73
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      52 KIRYKSKRQTSGGPVDSSEYQOELRELFKLQMGFNADMTPTFKFDPKFEVIEK 111

QY      74 POS 76
      ||:
Db      112 PQA 114

RESULT 11
US-09-205-679-1
; Sequence 1, Application US/09205679
; Patent No. 6048718
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: NOVEL ATP SYNTHASE COUPLING FACTOR 6
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PasterSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/205,679
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/828,239
; FILING DATE: 03/31/1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
```

```

;
; REFERENCE/DOCKET NUMBER: PF-0260 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: 2496341
;
US-09-205-679-1

Query Match      62.7%; Score 252; DB 3; Length 114;
Best Local Similarity 71.4%; Pred. No. 1.9e-24;
Matches 45; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY      14 KIRRYKAKRLASGGPVDTGPTGYQGVDRBELFKLQMTYKSGMDKFPPTNFEDPKFEVIDK 73
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      52 KIRYKSKRQTSGGPVDSSEYQOELRELFKLQMGFNADMTPTFKFDPKFEVIEK 111

QY      74 POS 76
      ||:
Db      112 PQA 114

RESULT 12
US-09-513-999C-5968
; Sequence 5968, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J. B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J. Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; PATENT NO. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5968
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 65
; OTHER INFORMATION: Xaa-Pro or Ser
;
US-09-513-999C-5968

Query Match      38.8%; Score 156; DB 4; Length 69;
Best Local Similarity 78.4%; Pred. No. 1.6e-12;
Matches 29; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      1 NKEIDPVQKLFVDKIRRYKAKRLASGGPVDTGPTGYQ 37
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      33 NKEIDPQKLFVDKIRYKSKRQTSGGPVDSXSYQ 69

RESULT 13
US-09-489-039A-7653
; Sequence 7653, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLBSTIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
```

```

1 / FILE REFERENCE: 2709.2004001 /
2 / CURRENT APPLICATION NUMBER: US/09/489, 0394
3 / CURRENT FILING DATE: 2000-01-27
4 / PRIORITY APPLICATION NUMBER: US 60/117,747
5 / PRIORITY FILING DATE: 1999-01-29
6 / NUMBER OF SEQ ID NOS: 1342
7 / SEQ ID NO 7653
8 / LENGTH: 583
9 / TYPE: PRT
10 / ORGANISM: Klebsiella pneumoniae
11 / US-09-489-039A-7653

```

Query Match	16.7%	Score 67	DB 4	Length 583
Best Local Similarity	24.4%	Pred. No. 4.8		
Matches 22	Conservative 10	Mismatches 22	Indels 36	Gaps 3

[illegible]

```

RESULT 14
US-09-538-092-901
Sequence 901, Application US/09538092
Patent No. 6753314
GENERAL INFORMATION:
APPLICANT: Glot, Loic
APPLICANT: Mansfield, Traci A.
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
FILE REFERENCE: 15966-542
CURRENT APPLICATION NUMBER: US/09/538,092
CURRENT FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/1127,352
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/178,965
PRIOR FILING DATE: 2000-02-01
NUMBER OF SEQ ID NOS: 1387
SOFTWARE: CuraPatSeqFormatter Version 0.9
SEQ ID NO 901
LENGTH: 1940
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (0)...(0)
OTHER INFORMATION: Polypeptide Accession Number P11055
US-09-538-092-901

```

Query Match 16.0%; Score 64.5; DB 4; Length 1940;  
Best Local Similarity 28.8%; Pred. No. 44;  
Matches 19; Conservative 12; Mismatches 34; Indels 1; Gaps 1.

QY 4LDEVOQKLFUDKIREYKAKRLASGPEVDTPBEYOQEVDRLEFKLKQMYGGMDEPPTNF 63  
 Db 1443 LDKKORNPDKVLEWFKTKKEESQALFASLSESLSTELFKLNAVEER-LDQETVVR 1501

Qy	64	EDPKFE	69
		:	
Db	1502	ENKNLE	1507

RESULT 15  
US-09-949-016-8888  
; Sequence 8888, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

```

1 FILE REFERENCE: CLO01307
2
3 CURRENT APPLICATION NUMBER: US/09/949,016
4
5 CURRENT FILING DATE: 2000-04-14
6
7 PRIOR APPLICATION NUMBER: 60/241,755
8
9 PRIOR FILING DATE: 2000-10-20
10
11 PRIOR APPLICATION NUMBER: 60/237,768
12
13 PRIOR FILING DATE: 2000-10-03
14
15 PRIOR APPLICATION NUMBER: 60/231,498
16
17 PRIOR FILING DATE: 2000-09-08
18
19 NUMBER OF SEQ ID NOS: 207012
20
21 SOFTWARE: FastSeq for Windows Version 4.0.
22
23 SEQ ID NO 8888
24
25 LENGTH: 1963
26
27 TYPE: PRT
28
29 ORGANISM: Human
30
31 US-09-949-016-8888

```

Query Match	16.0%;	Score	64.5;	DB	4;	Length	1663;
Best local Similarity	28.8%;	Pred. No.	45;				
Matches	19;	Conservative	12;	Mismatches	34;	Indels	1;
						Gaps	1;

QY 4 LDPVQKFLDKIREYKARLASSGPPVITGPYEQOEVRRLFTXKQMYGGMGEMKFPTEPNF 63  
Db 1466 LDKQRNFDKVLAEWKTCEESDAELIASLKESRSLSTELFKLKNAYEEA-LDQLETVYKR 1524

QY	64	EDPKFE	69
		:	
Db	1525	ENKNLE	1530

```
Search completed: April  4, 2005, 18:59:21
Job time : 31 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: April 4, 2005, 18:55:50 ; Search time 94.5 Seconds  
(without alignments)  
266.678 Million cell updates/sec

Sequence: 1 NKEIDPVQKLFIDKIREYKA.....KPTFNFEDPKFEVLDPQS 76

Searched: 1413372 beqs, 331592847 residues

```
Total number of hits satisfying chosen parameters: 1413372
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%
	Listing first	45 summaries

Database : Published Applications\_AA:\*

1:	/cgm2_6/prodataa/1/pubpaaf/US07_PUBCOMB.pcp.*
2:	/cgm2_6/prodataa/1/pubpaaf/PCST_NEW_PUB.pcp.*
3:	/cgm2_6/prodataa/1/pubpaaf/US06_NEW_PUB.pcp.*
4:	/cgm2_6/prodataa/1/pubpaaf/US06_PUBCOMB.pcp.*
5:	/cgm2_6/prodataa/1/pubpaaf/US07_NEW_PUB.pcp.*
6:	/cgm2_6/prodataa/1/pubpaaf/PCrUS_PUBCOMB.pcp.*
7:	/cgm2_6/prodataa/1/pubpaaf/US08_NEW_PUB.pcp.*
8:	/cgm2_6/prodataa/1/pubpaaf/US08_PUBCOMB.pcp.*
9:	/cgm2_6/prodataa/1/pubpaaf/US09A_PUBCOMB.pcp.*
10:	/cgm2_6/prodataa/1/pubpaaf/US09B_PUBCOMB.pcp.*
11:	/cgm2_6/prodataa/1/pubpaaf/US09C_PUBCOMB.pcp.*
12:	/cgm2_6/prodataa/1/pubpaaf/US09_PUBCOMB.pcp.*
13:	/cgm2_6/prodataa/1/pubpaaf/US10A_PUBCOMB.pcp.*
14:	/cgm2_6/prodataa/1/pubpaaf/US10C_PUBCOMB.pcp.*
15:	/cgm2_6/prodataa/1/pubpaaf/US10D_PUBCOMB.pcp.*
16:	/cgm2_6/prodataa/1/pubpaaf/US10E_PUBCOMB.pcp.*
17:	/cgm2_6/prodataa/1/pubpaaf/US10_NEW_PUB.pcp.*
18:	/cgm2_6/prodataa/1/pubpaaf/US11_NEW_PUB.pcp.*
19:	/cgm2_6/prodataa/1/pubpaaf/US60_NEW_PUB.pcp.*
20:	/cgm2_6/prodataa/1/pubpaaf/US60_PUBCOMB.pcp.*

**Pred.** No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	358	89.1	108	17	US-10-92e-543-77	Sequence 77, Apr
2	315	78.4	84	14	US-10-10e-698-6594	Sequence 6594, Apr
3	315	78.4	108	15	US-10-131-410-119	Sequence 119, Apr
4	315	78.4	108	16	US-10-40f-765A-1253	Sequence 1253, Apr
5	105	26.1	56	9	US-09-864-761-33712	Sequence 33712, Apr
6	105	26.1	56	9	US-09-864-761-33712	Sequence 33712, Apr
7	72.5	18.0	858	15	US-10-369-493-5536	Sequence 5536, Apr
8	72.5	18.0	858	15	US-10-369-493-5537	Sequence 5537, Apr
9	71.5	17.8	440	15	US-10-369-493-13341	Sequence 13341, Apr
10	70.5	17.5	262	14	US-10-156-761-11145	Sequence 11145, Apr
11	66.5	16.5	103	16	US-10-437-963-140910	Sequence 140910, Apr
12	66.5	16.5	219	15	US-10-29f-115-1011	Sequence 1011, Apr
13	66.5	16.5	458	14	US-10-21b-181-14	Sequence 14, Apr

14	66.5	16.5	458	14	US-10-212-912-12.14	Sequence 14, Appl
15	66.5	16.5	458	14	US-10-213-044-14	Sequence 14, Appl
16	66.5	16.5	458	14	US-10-213-182-14	Sequence 14, Appl
17	66.5	16.5	458	14	US-10-213-060-14	Sequence 14, Appl
18	66.5	16.5	458	14	US-10-213-052-14	Sequence 4, Appl
19	66.5	16.5	489	14	US-10-273-517-4	Sequence 4, Appl
20	66.5	16.5	489	15	US-10-311-104-4	Sequence 4, Appl
21	65	16.2	443	15	US-10-335-977-7965	Sequence 7965, Appl
22	65	16.2	448	15	US-10-335-977-7966	Sequence 7966, Appl
23	64.5	16.0	405	15	US-10-424-599-174941	Sequence 147941, Appl
24	64.5	16.0	1940	16	US-09-738-630-0.99	Sequence 99, Appl
25	64.5	16.0	1940	16	US-10-408-7658-11.75	Sequence 1175, Appl
26	64	15.9	316	13	US-10-001-870-10.13	Sequence 13, Appl
27	64	15.9	320	15	US-10-424-599-256965	Sequence 256965, Appl
28	64	15.9	859	15	US-10-282-1224-65095	Sequence 65095, Appl
29	63.5	15.8	288	14	US-10-424-599-9275777	Sequence 275777, Appl
30	63.5	15.8	634	14	US-10-153-668-100	Sequence 100, Appl
31	63.5	15.8	1118	14	US-10-153-668-104	Sequence 104, Appl
32	63.5	15.8	1117	15	US-10-336-472-16	Sequence 16, Appl
33	63.5	15.8	1859	15	US-10-336-472-20	Sequence 20, Appl
34	63.5	15.8	1935	15	US-10-336-472-32	Sequence 32, Appl
35	63	15.7	458	15	US-10-369-493-3701	Sequence 3701, Appl
36	63	15.7	674	16	US-10-437-963-110667	Sequence 110667, Appl
37	63	15.7	782	16	US-10-437-963-188046	Sequence 188046, Appl
38	62.5	15.5	263	14	US-10-029-386-33579	Sequence 33579, Appl
39	62.5	15.5	485	14	US-10-408-7658-443	Sequence 443, Appl
40	62.5	15.5	794	14	US-10-128-714-3165	Sequence 3165, Appl
41	62.5	15.5	1138	14	US-10-128-714-8165	Sequence 8165, Appl
42	62.5	15.5	1185	14	US-10-320-797-3350	Sequence 3350, Appl
43	62	15.4	245	15	US-10-425-114-59597	Sequence 59597, Appl
44	62	15.4	259	15	US-10-369-493-6882	Sequence 6882, Appl
45	62	15.4	498	15	US-10-424-599-240937	Sequence 240937, Appl

## ALIGNMENTS

```

US-10-926-543-77
RESULT 1
US-10-926-543-77
; Sequence 77, Application US/10926543
; Publication No. US2005004859A1
; GENERAL INFORMATION:
; APPLICANT: Jendoubi, Moncef
; TITLE OF INVENTION: LUNG CANCER SPECIFIC
; TITLE OF INVENTION: ANTIBODIES AND THEIR
; TITLE OF INVENTION: MANAGEMENT OF LUNG C
; FILE REFERENCE: 705403.4004
; CURRENT APPLICATION NUMBER: US/10/926,544
; CURRENT FILING DATE: 2004-08-25
; PRIOR APPLICATION NUMBER: US 60/497,790
; PRIOR FILING DATE: 2003-08-25
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 77
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-926-543-77

```

Query Match	Score	DB	Length
89.1%	358	17	108

Matches 65; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

```
QY      1 NKELDPVQKLF.LDKIREYKAKRLASGSPVDTPGEYQOEVDRELFLKQMYGKGEMDKFPT 600
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      33 NKELDPVQKLF.VDKIREYKSKRQASGSPVDIGEPYQGDREL.YKLQMYGKGEMDTPPT 920
```

```
QY      61 FNFEDPKFEVLDPQS 76
          | : ||||| : |||||
Db      93 FKFDPPKFEVIDKPQS 108
```

## RESULT 2



```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33712
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000087.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
; OTHER INFORMATION: SWISSPROT HIT: P18859, EVALUATE 5.00e-24
; OTHER INFORMATION: EST_HUMAN HIT: AA128761.1, EVALUATE 5.00e-24
US-09-864-761-33712

Query Match      26.1%; Score 105; DB 9; Length 56;
Best Local Similarity 86.4%; Pred. No. 5.1e-05;
Matches 19; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Cy      1 NKELDPVOKLFVDKIREYKSR 22
Db      35 NKELDPIOKLFVDKIREYKSR 56

RESULT 6
US-09-864-761-34125
; Sequence 34125, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aegm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263, 6
; PRIOR FILING DATE: 2000-10-04
; CURRENT APPLICATION NUMBER: US 60/236,359
```

```

; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34125
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000139.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 5.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EST_HUMAN HIT: AA128761.1, EVALUATE 5.00e-24
; OTHER INFORMATION: SWISSPROT HIT: P18859, EVALUATE 5.00e-24
US-09-864-761-34125

Query Match      26.1%; Score 105; DB 9; Length 56;
Best Local Similarity 86.4%; Pred. No. 5.1e-05;
Matches 19; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Cy      1 NKELDPVOKLFVDKIREYKSR 22
Db      35 NKELDPIOKLFVDKIREYKSR 56

RESULT 7
US-10-369-493-5536
; Sequence 5536, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
```

```

; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5536
; LENGTH: 858
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5536

```

```

Query Match      18.0%; Score 72.5; DB 15; Length 858;
Best Local Similarity 35.8%; Pred. No. 12;
Matches 24; Conservative 8; Mismatches 26; Indels 9; Gaps 3;

```

```

QY 4 LDPVQKFLDKIRREYKAKRLASGGPVDGTPGEYQOEVDRELFLKQKMYGKGEMDKPPTFN 62
DB 399 LEETTRIFKIDRIPEKARADLKSCKKIVVDGKEYD-----YALGVNFGKGEPPDKMPFN 451
QY 63 F-EDPKF 68
DB 452 VNESPQF 458

```

## RESULT 8

```

US-10-369-493-5537
; Sequence 5537, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-101520521B
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5537
; LENGTH: 858
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5537

```

```

Query Match      18.0%; Score 72.5; DB 15; Length 858;
Best Local Similarity 35.8%; Pred. No. 12;
Matches 24; Conservative 8; Mismatches 26; Indels 9; Gaps 3;

```

```

QY 4 LDPVQKFLDKIRREYKAKRLASGGPVDGTPGEYQOEVDRELFLKQKMYGKGEMDKPPTFN 62
DB 399 LEETTRIFKIDRIPEKARADLKSCKKIVVDGKEYD-----YALGVNFGKGEPPDKMPFN 451

```

```

QY 63 F-EDPKF 68
DB 452 VNESPQF 458

```

## RESULT 9

```

US-10-369-493-13341
; Sequence 13341, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-101520521B
; CURRENT APPLICATION NUMBER: US/10369,493

```

```

; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13341
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(440)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-13341

```

```

Query Match      17.8%; Score 71.5; DB 15; Length 440;
Best Local Similarity 39.0%; Pred. No. 6.7;
Matches 16; Conservative 9; Mismatches 13; Indels 3; Gaps 1;
QY 7 VQKFLDKIRREYKAKRLASGGPVDGTPGEYQOEVDRELFL 44
DB 132 VQQLFVDLVKYLDTSAVRVVTGGPAETGCLQKRFDTIFF 172

```

## RESULT 10

```

US-10-156-761-11145
; Sequence 11145, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11145
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11145

```

```

Query Match      17.5%; Score 70.5; DB 14; Length 262;
Best Local Similarity 30.1%; Pred. No. 4.7;
Matches 25; Conservative 11; Mismatches 34; Indels 13; Gaps 2;

```

```

QY 5 DPVQKFLDKIRREYKAKRLASGGPVDGTPGEYQOEVDRELFLKQKMYGKGEMDKPPTFN 64
DB 100 DPDDQLFMPTVEDVAFGPAAGL--TGALELARVDRLADQVGAFFKGRPPHHLSFGOR 157
QY 65 -----DPKEVDLKPQS 76
DB 158 RRVAATVLAAMEPEILVLDEPSS 180

```

## RESULT 11

```

US-10-437-963-140910
; Sequence 140910, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei

```



```
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT FILING DATE: 2003-05-14
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 140910
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_42063C.1.pep
US-10-437-963-140910

Query Match
Best Local Similarity 30.0%; Pred. No. 4.4;
Matches 18; Conservative 12; Mismatches 23; Indels 7; Gaps 2;

Qy 6 PVQKLFIDKIRRYKAKRLASGSPVDTGPEYQOE-----FKLKQMTGKGMKFFTFN 62
Db 31 PLQNVFSPKVFYKAFAPNFYFGNLMEVPGYKKGVGKIKLRPFKIK---KGVQGFPPPLN 86

RESULT 12
US-10-296-115-1011
; Sequence 1011, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296.115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 1011
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-296-115-1011

Query Match
Best Local Similarity 16.5%; Score 66.5; DB 15; Length 219;
Matches 25; Conservative 10; Mismatches 32; Indels 39; Gaps 3;

Qy 4 LDPVQKLFIDKIRRYKAKRLASGSPVDTGPEYQOE-----VDRELFLKLMQYKKG 53
Db 34 IDTWAMFTEER--DLDLVTLVFGSPDSTGHRYPSPRRRMVQVDRITGVYLRSTARN 91

Qy 54 -----EMDKPFTNFEDPKFEVLD 72
Db 92 HLTDRNLNLTSDHGMTTVDKRAQDLVEFHKFPNFTFRDIFELLD 137

RESULT 13
US-10-213-181-14
; Sequence 14, Application US/10213181
; Publication No. US20030054484A1
; GENERAL INFORMATION:
; APPLICANT: Fong, Sherman
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul
; APPLICANT: Grimaldi, Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Hillan, Kenneth
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin
```

```
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Immune
; TITLE OF INVENTION: Related Diseases
; FILE REFERENCE: P133R1C7
; CURRENT APPLICATION NUMBER: US/10/213.181
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 10/052,594
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: PCT/US00/30873
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: US 60/177,118
; PRIOR FILING DATE: 2000-01-20
; NUMBER OF SEQ ID NOS: 24
; SEQ ID NO 14
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-213-181-14

Query Match
Best Local Similarity 16.5%; Score 66.5; DB 14; Length 458;
Matches 25; Conservative 10; Mismatches 32; Indels 39; Gaps 3;

Qy 4 LDPVQKLFIDKIRRYKAKRLASGSPVDTGPEYQOE-----VDRELFLKLMQYKKG 53
Db 176 IDTWAMFTEER--DLDLVTLVFGSPDSTGHRYPSPRRRMVQVDRITGVYLRSTARN 233

Qy 54 -----EMDKPFTNFEDPKFEVLD 72
Db 234 HLTDRNLNLTSDHGMTTVDKRAQDLVEFHKFPNFTFRDIFELLD 279

RESULT 14
US-10-212-912-14
; Sequence 14, Application US/10212912
; Publication No. US2003007737A1
; GENERAL INFORMATION:
; APPLICANT: Fong, Sherman
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul
; APPLICANT: Grimaldi, Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Hillan, Kenneth
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Immune
; TITLE OF INVENTION: Related Diseases
; FILE REFERENCE: P133R1C2
; CURRENT APPLICATION NUMBER: US/10/212.912
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 10/052,594
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: PCT/US00/30873
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: US 60/172,059
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 24
; SEQ ID NO 14
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-212-912-14

Query Match
Best Local Similarity 16.5%; Score 66.5; DB 14; Length 458;
Matches 25; Conservative 10; Mismatches 32; Indels 39; Gaps 3;

Qy 4 LDPVQKLFIDKIRRYKAKRLASGSPVDTGPEYQOE-----VDRELFLKLMQYKKG 53
Db 176 IDTWAMFTEER--DLDLVTLVFGSPDSTGHRYPSPRRRMVQVDRITGVYLRSTARN 233
```

OY 54 -----EMDKPTNFEDPKPEYLD 72  
DB 234 HLTDRNLNLTSDHGMTTVDRKAGDLVEFHKFPNFTFRDIEFELD 279

RESULT 15  
US-10-213-044-14  
; Sequence 14, Application US/10213044  
; Publication No. US2003007738A1  
; GENERAL INFORMATION:  
; APPLICANT: Fong, Sherman  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul  
; APPLICANT: Grimaldi, Christopher  
; APPLICANT: Gurney, Austin  
; APPLICANT: Hillan, Kenneth  
; APPLICANT: Tumaas, Daniel  
; APPLICANT: Watanabe, Colin  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Immune  
; TITLE OF INVENTION: Related Diseases  
; FILE REFERENCE: P3133R1C6  
; CURRENT APPLICATION NUMBER: US/10/213,044  
; CURRENT FILING DATE: 2002-08-05  
; PRIOR APPLICATION NUMBER: US 10/052,594  
; PRIOR FILING DATE: 2002-01-18  
; PRIOR APPLICATION NUMBER: PCT/US00/30873  
; PRIOR FILING DATE: 2000-11-10  
; PRIOR APPLICATION NUMBER: US 60/177,118  
; PRIOR FILING DATE: 2000-01-20  
; NUMBER OF SEQ ID NOS: 24  
; SEQ ID NO 14  
; LENGTH: 458  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-213-044-14

Query Match 16.5%; Score 66.5; DB 14; Length 458;  
Best Local Similarity 23.6%; Pred. No. 28;  
Matches 25; Conservative 10; Mismatches 32; Indels 39; Gaps 3;  
OY 4 LDPVQKFLDKIREYKARLASGSPVDTGPEYQOE-----VDRELFLKQMYGKG 53  
DB 176 IDTWAMFTBE--DLDLVTLVYFGEPDSTGHRYPGSPERRRMVQVDRTVGYLRESIARN 233  
OY 54 -----EMDKPTNFEDPKPEYLD 72  
DB 234 HLTDRNLNLTSDHGMTTVDRKAGDLVEFHKFPNFTFRDIEFELD 279

Search completed: April 4, 2005, 19:15:11  
Job time : 103.5 sec

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2005, 18:58:20 ; Search time 124.5 Seconds

(without alignments)  
236.095 Million cell updates/sec

Title: US-09-831-951A-2

Perfect score: 402  
Sequence: 1 NKEADPVQKFLDKIREYKA.....KPTFNPEDPKFEVLDKRQS 76

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*\n2: Geneseqp1990s:\*\n3: Geneseqp2000s:\*\n4: Geneseqp2001s:\*\n5: Geneseqp2002s:\*\n6: Geneseqp2003as:\*\n7: Geneseqp2003bs:\*\n8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	402	100.0	76	4	AAB80964 Rat Coupl
2	402	100.0	108	7	AD663636 Rat Prote
3	402	100.0	108	7	AD663632 Rat Prote
4	402	100.0	108	7	AD663628 Rat Prote
5	402	100.0	108	7	AD663624 Rat Prote
6	317	78.9	76	4	AAB80963 Human Cou
7	315	78.4	84	4	AAQ75820 Human col
8	315	78.4	108	2	AAV48588 Human bre
9	315	78.4	108	3	AAQ04005 Human sec
10	315	78.4	108	7	AD663626 Human Pro
11	315	78.4	108	7	AD663638 Human Pro
12	315	78.4	108	7	AD663634 Human Pro
13	315	78.4	108	7	AD663630 Human Pro
14	315	78.4	108	7	ADJ69447 Human hea
15	309	76.9	106	3	AAAG01886 Human sec
16	252	62.7	114	2	AAW87500 Human mit
17	252	62.7	114	3	AAV82605 Human mit
18	202	50.2	39	4	AAB80972 Rat CF6 P
19	156.5	38.9	106	3	ABB60627 Drosophil
20	156	38.8	69	3	AAAG01887 Human sec
21	132.5	33.0	147	4	ABB57878 Drosophil
22	105	26.1	20	4	AAB80968 Rat CF6 P
23	105	26.1	20	4	AAB80969 Rat CF6 P
24	105	26.1	56	4	AAAM1390 Peptide #
25	105	26.1	56	4	AAAM14420 Peptide #

26	105	26.1	56	4	ABB33368 Peptide #
27	105	26.1	56	4	ABB32935 Peptide #
28	105	26.1	56	4	AAAM26396 Peptide #
29	105	26.1	56	4	AAAM26833 Peptide #
30	105	26.1	56	4	ABB27764 Human pep
31	105	26.1	56	4	ABB28193 Human pep
32	105	26.1	56	4	ABB18414 Protein #
33	105	26.1	56	4	ABB18827 Protein #
34	105	26.1	56	4	AAAM6547 Human bon
35	105	26.1	56	4	AAAM66119 Human bon
36	105	26.1	56	4	AAAM53736 Human bra
37	105	26.1	56	4	AAAM54153 Human bra
38	105	26.1	56	4	ABG47787 Human liv
39	105	26.1	56	4	ABG48215 Human liv
40	105	26.1	56	4	AAAM02147 Peptide #
41	105	26.1	56	4	AAAM01731 Peptide #
42	105	26.1	56	5	ABG35769 Human pep
43	105	26.1	56	5	ABG36199 Human pep
44	84	20.9	16	4	AAB80973 Rat CF6 P
45	84	20.9	18	2	ADO24846 Human EBZ

## ALIGNMENTS

RESULT 1  
AAB80964  
ID AAB80964 standard; protein: 76 AA.  
XX  
AC AAB80964;  
XX  
DT 08-JUN-2001 (first entry)  
XX  
DE Rat Coupling factor 6 CF6.  
XX  
KW Rat; coupling factor 6; CF6; antiinflammatory; antidiabetic; antitumor;  
KW cardiac; hypotensive; antiangiogenic; antiaesthetic; antineurotic;  
KW antiarthritic; proton-translocating ATPase; prostaglandin;  
KW cytoplasmic PLA2; cardiovascular disorder; diabetes;  
KW inflammatory disorder; rheumatoid arthritis; cerebrovascular disorder.  
XX  
OS Rattus sp.  
XX  
PN WO200121205-A1.  
XX  
PD 29-MAR-2001.  
XX  
PF 03-AUG-2000; 2000WO-JP005210.  
XX  
PR 17-SEP-1999; 99JP-00264687.  
XX  
PA (SUNR ) SUNTORY LTD.  
XX  
PI Osanal T, Magoca K;  
XX  
WP: 2001-266047/27.  
XX  
PT Coupling factor 6 useful in diagnosis of diseases relating to excess or  
PT lack of prostaglandin and cytoplasmic PLA2 activity in blood, and in  
PT screening drugs for treating e.g. cardiovascular infection.  
XX  
PS Claim 15; Page 56-77; 70pp; Japanese.  
XX  
CC The present sequence is rat coupling factor 6 (CF6) protein. CF6 is a  
CC subunit of a proton-translocating ATPase found in mitochondria. CF6 can  
CC be used in diagnosis of diseases relating to excess or lack of  
CC prostaglandin and cytoplasmic PLA2 activity, and in screening inhibitors  
CC and potentiators as drugs to treat e.g. cardiovascular infection,  
CC hypertension, arteriosclerosis, angina pectoris, cardiac insufficiency,  
CC cerebrovascular disorders, hyperlipemia, diabetes, bronchitis, gastric  
CC ulcer, pregnant eclampsia, haemolytic uremia syndrome, thrombopenic  
CC purpura, inflammatory diseases like cerebral infarction, acute  
CC pancreatitis, asthma, ARDS (adult respiratory distress syndrome), and

CC rheumatoid arthritis  
XX Sequence 76 AA;  
SQ

Query Match 100.0%; Score 402; DB 4; Length 76;  
Best Local Similarity 100.0%; Pred. No. 2e-42;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKEIDPVQKFLDKIREYKAKRLASGGPVDTGPEYQOEVDRELFKLKQWYKGMKMDKFPPT 60  
DB 1 NKEIDPVQKFLDKIREYKAKRLASGGPVDTGPEYQOEVDRELFKLKQWYKGMKMDKFPPT 60

QY 61 FNFEDPKFEVLDPKQPS 76  
DB 61 FNFEDPKFEVLDPKQPS 76

RESULT 2  
ID ADE63636 standard; protein; 108 AA.  
XX ADE63636;  
AC  
XX 29-JAN-2004 (first entry)  
DT  
XX  
DE Rat Protein P21571, SEQ ID NO 9580.  
XX  
KM Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
XX  
OS Rattus norvegicus.  
XX  
PN W02003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002MO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
PA (GEHO) GEN HOSPITAL CORP.  
PA (FARB) BAYER AG.  
XX  
PI Woolf C, D'urso D, Befort K, Costigan M;  
XX  
DR WPI; 2003-268312/26.  
DR GENBANK; P21571.  
XX  
PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
PS Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more

CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 108 AA;  
XX

Query Match 100.0%; Score 402; DB 7; Length 108;  
Best Local Similarity 100.0%; Pred. No. 3.2e-42;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKEIDPVQKFLDKIREYKAKRLASGGPVDTGPEYQOEVDRELFKLKQWYKGMKMDKFPPT 60  
DB 33 NKEIDPVQKFLDKIREYKAKRLASGGPVDTGPEYQOEVDRELFKLKQWYKGMKMDKFPPT 92

QY 61 FNFEDPKFEVLDPKQPS 76  
DB 93 FNFEDPKFEVLDPKQPS 108

RESULT 3  
ID ADE63632 standard; protein; 108 AA.  
XX ADE63632;  
AC  
XX 29-JAN-2004 (first entry)  
DT  
XX  
DE Rat Protein P21571, SEQ ID NO 9576.  
XX  
KM Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
XX  
OS Rattus norvegicus.  
XX  
PN W02003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002MO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
PA (GEHO) GEN HOSPITAL CORP.  
PA (FARB) BAYER AG.  
XX  
PI Woolf C, D'urso D, Befort K, Costigan M;  
XX  
DR WPI; 2003-268312/26.  
DR GENBANK; P21571.  
XX  
PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
PS Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates

CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 108 AA;  
SQ

Query Match 100.0%; Score 402; DB 7; Length 108;  
Best Local Similarity 100.0%; Pred. No. 3.2e-42;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKEIDPVQKFLDKIREYKAKRLASGGPVDTGPEYQOEVDRELFLTKOMYKGEMDKRPT 60  
DB 33 NKEIDPVQKFLDKIREYKAKRLASGGPVDTGPEYQOEVDRELFLTKOMYKGEMDKRPT 92  
DB 61 FNFEDEPKREVLDPKPOS 76  
QY 93 FNFEDEPKREVLDPKPOS 108

Db 61 FNFEDEPKREVLDPKPOS 76  
QY 93 FNFEDEPKREVLDPKPOS 108

RESULT 4  
ADE63628  
ID ADE63628 standard; protein; 108 AA.  
XX  
AC ADE63628;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Rat Protein P21571, SEQ ID NO 9572.  
XX  
KM Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
KM chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
XX  
OS Rattus norvegicus.  
XX  
PN WO2003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
PA (GEHO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
PI Woolf C, D'urso D, Befort K, Costigan M;  
XX  
DR WPI; 2003-268312/26.  
DR GENBANK; P21571.  
XX  
PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
PS Claim 1; Page; 1017pp; English.  
XX  
CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,

CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 108 AA;  
SQ

Query Match 100.0%; Score 402; DB 7; Length 108;  
Best Local Similarity 100.0%; Pred. No. 3.2e-42;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKEIDPVQKFLDKIREYKAKRLASGGPVDTGPEYQOEVDRELFLTKOMYKGEMDKRPT 60  
DB 33 NKEIDPVQKFLDKIREYKAKRLASGGPVDTGPEYQOEVDRELFLTKOMYKGEMDKRPT 92  
DB 61 FNFEDEPKREVLDPKPOS 76  
QY 93 FNFEDEPKREVLDPKPOS 108

Db 61 FNFEDEPKREVLDPKPOS 76  
QY 93 FNFEDEPKREVLDPKPOS 108

RESULT 5  
ADE63624  
ID ADE63624 standard; protein; 108 AA.  
XX  
AC ADE63624;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Rat Protein P21571, SEQ ID NO 9568.  
XX  
KM Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
KM chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
XX  
OS Rattus norvegicus.  
XX  
PN WO2003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
PA (GEHO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
PI Woolf C, D'urso D, Befort K, Costigan M;  
XX  
DR WPI; 2003-268312/26.  
DR GENBANK; P21571.

XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX

PS Claim 1; Page: 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene  
CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 108 AA;

SO Query Match 100.0%; Score 402; DB 7; Length 108;  
Best Local Similarity 100.0%; Pred. No. 3.2e-42;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKELDPPVQKFLDKIRRYKAKRLASGSPVDTGPEYQOEVDRELFLKQMTGKGMKPEPT 60  
DB 33 NKELDPPVQKFLDKIRRYKAKRLASGSPVDTGPEYQOEVDRELFLKQMTGKGMKPEPT 92  
QY 61 FNFEDEPKFEVLDPKPOS 76  
DB 93 FNFEDEPKFEVLDPKPOS 108

RESULT 6  
AAB80963  
ID AAB80963 standard; protein; 76 AA.

AC AAB80963;

DT 08-JUN-2001 (first entry)

DE Human Coupling factor 6 CF6.

XX Human; coupling factor 6; CF6; antiinflammatory; antidiabetic; antilucer;  
KM cardiant; hypotensive; antiasthmatic; antirheumatic;  
KM antiarthritic; proton-transporting ATPase; prostaglandin;  
KM cytoplasmic PLA 2; cardiovascular disorder; diabetes;  
KM inflammatory disorder; rheumatoid arthritis; cerebrovascular disorder.

OS Homo sapiens.

XX WO200121205-A1.

XX 29-MAR-2001.

PF 03-AUG-2000; 2000WO-JP005210.

PR 17-SEP-1999; 99JP-00264687.

XX (SUNR) SUNTORY LTD.

XX Osana T, Magoca K;

XX WPI; 2001-266047/27.

PT Coupling factor 6 useful in diagnosis of diseases relating to excess or  
PT lack of prostaglandin and cytoplasmic PLA2 activity in blood, and in  
PT screening drugs for treating e.g. cardiovascular infarction.

PS Claim 15; Page 56; 70pp; Japanese.

XX The present sequence is human coupling factor 6 (CF6) protein. CF6 is a  
CC subunit of a proton-transporting ATPase found in mitochondria. CF6 can  
CC be used in diagnosis of diseases relating to excess or lack of  
CC prostaglandin and cytoplasmic PLA 2 activity, and in screening inhibitors  
CC and potentiators as drugs to treat e.g. cardiovascular infarction,  
CC hypertension, arteriosclerosis, angina pectoris, cardiac insufficiency,  
CC cerebrovascular disorders, hyperlipemia, diabetes, bronchitis, gastric  
CC ulcer, pregnant eclampsia, haemolytic uremia syndrome, thrombopenic  
CC purpura, inflammatory diseases like cerebral infarction, acute  
CC pancreatitis, asthma, ARDS (adult respiratory distress syndrome), and  
CC rheumatoid arthritis

XX Sequence 76 AA;

SO Query Match 78.9%; Score 317; DB 4; Length 76;  
Best Local Similarity 75.0%; Pred. No. 9.3e-32;  
Matches 57; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

QY 1 NKELDPPVQKFLDKIRRYKAKRLASGSPVDTGPEYQOEVDRELFLKQMTGKGMKPEPT 60  
DB 1 NKELDPPVQKFLDKIRRYKAKRLASGSPVDTGPEYQOEVDRELFLKQMTGKGMKPEPT 60  
QY 61 FNFEDEPKFEVLDPKPOS 76  
DB 61 FNFEDEPKFEVLDPKPOS 76

RESULT 7  
AAG75820  
ID AAG75820 standard; protein; 84 AA.

AC AAG75820;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen protein SEQ ID NO:6584.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;  
XX colorectal carcinoma; chromosome 10.

OS Homo sapiens.

XX WO200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US026524.

XX 29-SEP-1999; 99US-0157137P.

XX 03-NOV-1999; 99US-0163280P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI; 2001-235357/24.

XX N-PSDB; AAH35225.

PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,



CC and to design expression and secretion vectors  
XX  
SQ Sequence 108 AA;  
Query Match 78.4%; Score 315; DB 3; Length 108;  
Best Local Similarity 73.7%; Pred. No. 2.6e-31;  
Matches 56; Conservative 11; Mismatches 9; Indels 0; Gaps 0;  
OY 1 NKEIDPVQKLFDRKREYKAKRLASGSPVDTGPEYQOEVDRELFLKQMGKGMDFPT 60  
DB 33 NKEIDPQKLFVFKIREYKSKRQTSQGPVDASSEYQOELERELFLKQMGFNADMTFT 92  
OY 61 FNFEDPKFEVLDKPKQS 76  
DB 93 FKFEDEPKFEVIEKPKQA 108  
RESULT 10  
ADE63626  
ID ADE63626 standard; protein; 108 AA.  
AC ADE63626;  
XX  
XX 29-JAN-2004 (first entry)  
DT  
XX Human Protein P18859, SEQ ID NO 9570.  
DE  
XX Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KM spared nerve injury; SNI; Chung.  
XX Homo sapiens.  
OS  
XX WO2003016475-A2.  
PN  
XX 27-FEB-2003.  
PD  
XX 14-AUG-2002; 2002WO-US025765.  
PF  
XX 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
XX (GHEO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
XX Woolf C, D'urso D, Befort K, Costigan M;  
PI WPI; 2003-268312/26.  
DR GENBANK; P18859.  
DR  
XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
XX Claim 1; Page; 1017pp; English.  
PS  
XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating

CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 108 AA;  
Query Match 78.4%; Score 315; DB 7; Length 108;  
Best Local Similarity 73.7%; Pred. No. 2.6e-31;  
Matches 56; Conservative 11; Mismatches 9; Indels 0; Gaps 0;  
OY 1 NKEIDPVQKLFDRKREYKAKRLASGSPVDTGPEYQOEVDRELFLKQMGKGMDFPT 60  
DB 33 NKEIDPQKLFVFKIREYKSKRQTSQGPVDASSEYQOELERELFLKQMGFNADMTFT 92  
OY 61 FNFEDPKFEVLDKPKQS 76  
DB 93 FKFEDEPKFEVIEKPKQA 108  
RESULT 11  
ADE63638  
ID ADE63638 standard; protein; 108 AA.  
XX  
XX ADE63638;  
AC  
XX 29-JAN-2004 (first entry)  
DT  
XX Human Protein P18859, SEQ ID NO 9582.  
DE  
XX Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KM spared nerve injury; SNI; Chung.  
XX  
XX Homo sapiens.  
OS  
XX WO2003016475-A2.  
PN  
XX 27-FEB-2003.  
PD  
XX 14-AUG-2002; 2002WO-US025765.  
PF  
XX 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
XX (GHEO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
XX Woolf C, D'urso D, Befort K, Costigan M;  
PI WPI; 2003-268312/26.  
DR GENBANK; P18859.  
DR  
XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
XX Claim 1; Page; 1017pp; English.  
PS  
XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence



CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity for identifying a compound or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 108 AA;

Query Match 78.4%; Score 315; DB 7; Length 108;  
Best Local Similarity 73.7%; Pred. No. 2.6e-31;  
Matches 56; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

QY 1 NKEIDPVQKFLDKIRRYKAKRLASGGPVDTPGEYQOEVDRELFLKQMYGKGMDEKPT 60  
Db 33 NKEIDPIQKLFVDKIRRYKSKRQTSGGPVDASSRYQOELELFLKQMPGNADMTFTPT 92

QY 61 FNFDPKFEVLDKPKOS 76  
Db 93 FKFDPKFEVLEKPOA 108

RESULT 12

ADE63634 ID ADE63634 standard; protein; 108 AA.

AC ADE63634;

DT 29-JAN-2004 (first entry)

DE Human Protein P18859, SEQ ID NO 9578.

KW Human; pain; neuronal tissue; gene therapy;  
KM spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

OS Homo sapiens.

PN WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

PA (GEHO ) GEN HOSPITAL CORP.

PA (FARB ) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

DR WPI; 2003-268312/26.

DR GENBANK; P18859.

PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
PS Claim 1; Page; 1017p; English.

XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 108 AA;

Query Match 78.4%; Score 315; DB 7; Length 108;  
Best Local Similarity 73.7%; Pred. No. 2.6e-31;  
Matches 56; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

QY 1 NKEIDPVQKFLDKIRRYKAKRLASGGPVDTPGEYQOEVDRELFLKQMYGKGMDEKPT 60  
Db 33 NKEIDPIQKLFVDKIRRYKSKRQTSGGPVDASSRYQOELELFLKQMPGNADMTFTPT 92

QY 61 FNFDPKFEVLDKPKOS 76  
Db 93 FKFDPKFEVLEKPOA 108

RESULT 13

ADE63630 ID ADE63630 standard; protein; 108 AA.

AC ADE63630;

DT 29-JAN-2004 (first entry)

DE Human Protein P18859, SEQ ID NO 9574.

KW Human; pain; neuronal tissue; gene therapy;  
KM spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

PN WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

PA (GEHO ) GEN HOSPITAL CORP.

PA (FARB ) BAYER AG.



PD 06-SEP-2000.

XX 21-FEB-2000; 2000EP-00200610.

XX 26-FEB-1999; 99US-0122487P.

XX (GENSET ) GENSET.

XX Dumas Mline Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

DR N-PSDB; AAC01892.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX Claim 13; SEQ ID NO 5967; 71pp + Sequence listing; English.

CC The present sequence is a polypeptide encoded by one of a large number of  
CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were  
CC prepared from total human RNA or polyA+ RNAs derived from 30 different  
CC tissues. EST sequences usually correspond mainly to the 3' untranslated  
CC region (UTR) of the mRNA because they are often obtained from oligo-dT  
CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA  
CC sequences derived from the 5' ends of mRNAs and even in those cases where  
CC longer cDNA sequences have been obtained, the full 5' UTR is rarely  
CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can  
CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs  
CC are also used in diagnostic, forensic, gene therapy and chromosome  
CC mapping procedures. They are used to obtain upstream regulatory sequences  
CC and to design expression and secretion vectors

XX Sequence 106 AA;

Query Match 76.9%; Score 309; DB 3; Length 106;

Best Local Similarity 74.3%; Pred. NO. 1.4e-30;

Matches 55; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

Qy 1 NKEIDPVOKLFLDKIREYKAKLASGSPVDTGPEYQOEVDRELFTLKOMYKGEMDKFPT 60  
Db 33 NKEIDPLOKLFVDKIREYKSKRQTSQGPVDASSEYQOEIRERLFTLKOMFGNADMTPTPT 92

Qy 61 FNPEDPKFEVLDPK 74

Db 93 FKFEDEPKFEVLEKP 106

Search completed: April 4, 2005, 19:19:23  
Job time : 125.5 secs

**This Page Blank (usp10)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2005, 18:59:30 ; Search time 118.5 seconds  
(without alignments)  
328.422 Million cell updates/sec

Title: US-09-831-951A-1  
Perfect score: 394  
Sequence: 1 NKELDPIQLFVDKIREYKS.....TFPTKPEDPKFEVLKRPQA 76

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot\_03:.\*  
1: uniprot\_sprot:.\*  
2: uniprot\_trembl:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	392	99.5	108	1	ATPR_HUMAN
2	383	97.2	108	2	Q6N259
3	363	92.1	108	2	Q8SPH6
4	346	87.8	108	1	ATPR_BOVIN
5	342	86.8	76	1	ATPR_PIG
6	332	84.3	108	1	ATPR_MOUSE
7	317	80.5	108	1	ATPR_RAT
8	245	62.2	107	2	Q6RFT1
9	237	60.2	107	2	Q6RGS5
10	182	46.2	112	2	Q6NYF7
11	180	45.7	117	2	Q7Q7P4
12	154.5	39.2	106	1	ATPR_DROME
13	118	29.9	147	2	Q9VZ72
14	111	26.2	54	2	Q812D0
15	79.5	20.2	1934	2	Q75UE1
16	79.5	20.2	1940	1	MYH3_HUMAN
17	78.5	19.9	557	2	Q7M3B1
18	78	19.8	1103	2	Q64XS8
19	77.5	19.7	285	2	Q6R840
20	77.5	19.7	764	2	Q91411
21	76.5	19.4	891	2	Q90WH5
22	76.5	19.4	1931	2	Q910C5
23	76.5	19.4	1940	1	MYH3_RAT
24	75.5	19.2	372	2	Q6LW9
25	75.5	19.2	611	2	Q14905
26	75.5	19.2	1491	2	Q75UE0
27	75.5	19.2	1931	2	Q91973
28	75.5	19.2	1934	1	MYH7_HUMAN
29	75.5	19.2	1935	1	MYH7_MESAU
30	75.5	19.2	1935	2	Q8MJU9
31	75.5	19.2	1935	2	Q9BE39

32	75.5	19.2	1935	2	Q9GKR1	Q9GKR1 sus scrofa
33	75.5	19.2	1935	2	Q91283	Q91283 mus musculus
34	75.5	19.2	1938	1	MYH6_MOUSE	Q02566 mus musculus
35	75.5	19.2	1938	1	MYH6_RAT	P02563 rattus norv
36	75.5	19.2	1939	1	MYH6_HUMAN	P13533 homo sapien
37	75.5	19.2	1939	1	MYH6_MESAU	P02564 rattus norv
38	73.5	18.7	1935	1	MYH7_RAT	Q94691 mus musculus
39	72.5	18.4	276	2	Q9D6G1	Q99020 mus musculus
40	72.5	18.4	285	1	ROAA_MOUSE	Q9GXR0 rattus norv
41	72.5	18.4	285	2	Q9GXR0	Q920U8 rattus norv
42	72.5	18.4	298	2	Q920U8	Q80XR6 mus musculus
43	72.5	18.4	311	2	Q80XR6	Q88311 rattus norv
44	72.5	18.4	331	2	Q88311	Q9GXR1 rattus norv
45	72.5	18.4	332	2	Q9GXR1	

## ALIGNMENTS

RESULT 1	ATPR_HUMAN	STANDARD	PRT	108 AA.
ID	ATPR_HUMAN	STANDARD	PRT	108 AA.
AC	P18859			
DT	01-NOV-1990 (Rel. 16, Created)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	ATP synthase coupling factor 6, mitochondrial precursor (EC 3.6.3.14)			
DE	(F6).			
GN	Name=ATP5F; Synonyms=ATPM;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=9131516; PubMed=1825642; DOI=10.1016/0378-1119(91)90068-M;			
RA	Javed A.A., Ogata K., Sanadi D.R.;			
RT	"Human mitochondrial ATP synthase: cloning cDNA for the nuclear-			
RT	encoded precursor of coupling factor 6.";			
RL	Gene 97:307-310(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=9131516; PubMed=1830479;			
RA	Higuchi T., Tsurumi C., Kawamura Y., Tsujita H., Osaka F.;			
RT	Yoshinaka Y., Tani I., Tanaka K., Ichihara A.;			
RT	"Molecular cloning of cDNA for the import precursor of human coupling			
RL	factor 6 of H(+)-ATP synthase in mitochondria.";			
RN	Biochem. Biophys. Res. Commun. 178:793-799(1991).			
RP	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Kidney;			
RX	MEDLINE=21154917; PubMed=11230166; DOI=10.1101/gr.154701;			
RA	Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glasel S.;			
RA	Ansorge W., Boecker M., Bloecher H., Bauerachs S., Blum H.;			
RA	Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.;			
RA	Wauson H.-W., Ottenwelder B., Obermaier B., Tampe J., Heubner D.;			
RA	Mewes R., Korn B., Klein M., Poustka A.;			
RT	"Towards a catalog of human genes and proteins: sequencing and			
RT	analysis of 500 novel complete protein coding human cDNAs.";			
RL	Genome Res. 11:422-435(2001).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20285799; PubMed=10830953; DOI=10.1038/35012518;			
RA	Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.;			
RA	Park H.-S., Toyoda A., Ishii K., Tocioki Y., Choi D.-K., Soeda E.;			
RA	Ohts M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.;			
RA	Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.;			
RA	Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.;			
RA	Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.;			
RA	Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.;			
RA	Mitsushima S., Shimizu N., Nordstiek G., Horstischer K., Brandt P.;			
RA	Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.;			
RA	Ramser J., Beck A., Klages S., Hennig S., Rieseemann L., Dagand E.;			

RA Mehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,  
 RA Lehrach H., Reinhardt R., Yaspo M.-L.;  
 RT "The DNA sequence of human chromosome 21.";  
 RL Nature 405:311-319(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Eye;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marisina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek U.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [6]  
 RP SEQUENCE OF 33-43.  
 RC TISSUE=Liver;  
 RX MEDLINE=93162045; PubMed=1286669;  
 RA Hochreiter D.F., Frutiger S., Paquet N., Bairoch A., Ravier F.,  
 RA Pasquali C., Sanchez J.-C., Tisot J.-D., Bjelqvist B., Vargas R.,  
 RA Appel R.D., Hughes G.J.;  
 RT "Human liver protein map: a reference database established by  
 RT microsequencing and gel comparison.";  
 RL Electrophoresis 13:992-1001(1992).  
 CC -1- FUNCTION: This is one of the chains of the nonenzymatic component  
 CC (CF0) subunit of the mitochondrial ATPase complex. F6 seems to  
 CC be part of the stalk that links CF0 to CF1(1). Also involved in  
 CC the restoration of oligomycin-sensitive ATPase activity to  
 CC depleted F1-F0 complexes.  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +  
 CC H(+) (out).  
 CC -1- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic  
 CC core - and CF(0) - the membrane proton channel. CF(0) seems to  
 CC have nine subunits: a, b, c, d, e, f, g, F6 and 8 (or A6L).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M37104; AAA51807.1; -;  
 DR EMBL: M73031; AAA58630.1; -;  
 DR EMBL: AL110183; CAB53667.1; -;  
 DR EMBL: AF001694; -; NOT ANNOTATED\_CDS.  
 DR EMBL: BC001178; AAH01178.1; -;  
 DR PIR: J70563; J70563.  
 DR SWISS-2DPAGE: P18859; HUMAN.  
 DR Genew: HGNC:847; ATP5J.  
 DR H-InvDB: HIX0016040; -;  
 DR Reactome: P18859; -;  
 DR MIM: 603152; -;  
 DR GO: GO:0005743; C:mitochondrial inner membrane; TAS.  
 DR GO: GO:0005739; C:mitochondrion; TAS.  
 DR InterPro: IPR008387; ATP synth\_F6.  
 DR Pfam: PF05511; ATP-synt\_F6; 1.  
 KW CF(0); Direct protein sequencing; Hydrogen ion transport;

KW Mitochondrion; Transic peptide.  
 FT TRANSIT 1 32 Mitochondrion.  
 FT CHAIN 33 108 ATP synthase coupling factor 6.  
 FT CONFLICT 68 68 Q -> H (in Ref. 2).  
 SQ SEQUENCE 108 AA; 12587 MW; EDC1A14F01A10F17 CRC64;  
 Query Match 99.5%; Score 392; DB 1; Length 108;  
 Best Local Similarity 98.7%; Pred. No. 1,1e-31;  
 Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NKELDPIQKLVKDKIRKYSKRTSGSPYDASSEYQOELERELFKLKQFGNADNMTPT 60  
 DB 33 NKELDPIQKLVKDKIRKYSKRTSGSPYDASSEYQOELERELFKLKQFGNADNMTPT 92  
 QY 61 FKEDPKFEVLEKPA 76  
 DB 93 FKEDPKFEVLEKPA 108  
 RESULT 2  
 ID Q6NZ59 PRELIMINARY; PRT; 108 AA.  
 AC Q6NZ59;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE ATP synthase, H+ transporting, mitochondrial F0 complex, subunit F6,  
 DE isoform a.  
 GN Name=ATP5J;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marisina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek U.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butcherfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX Direct MGC Project;  
 RA Submitted (FEB-2004) for the EMBL/Genbank/DBJ databases.  
 DR EMBL: BC066310; AAH66310.1; -;  
 DR GO: GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.  
 DR GO: GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.  
 DR GO: GO:0046961; F:hydrogen-transporting ATPase activity, rota.; IEA.  
 DR GO: GO:0015986; P:ATP synthesis coupled proton transport; IEA.  
 DR InterPro: IPR008387; ATP synth\_F6.  
 DR Pfam: PF05511; ATP-synt\_F6; 1.  
 SO SEQUENCE 108 AA; 12627 MW; EDD2C0CF01A10F17 CRC64;  
 Query Match 97.2%; Score 383; DB 2; Length 108;  
 Best Local Similarity 97.4%; Pred. No. 8.9e-31;

```

Matches 74; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NKELDPIQKLFVDKIREYKSKRQTSGGPVDASSEYQDLERELFKLKMFGNADMTPT 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 33 NKELDPIQKLFVDKIREYKSKRQTSGGPVDASSEYQDLERELFKLKMFGNADMTPT 92
QY 61 FKPEDPKFEVLEKPOA 76
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 93 FKPEDPKFEVLEKPOA 108

RESULT 3
Q8SPH6 PRELIMINARY; PRT; 108 AA.
ID Q8SPH6
AC Q8SPH6;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE ATP synthase subunit F6.
GN Name=ATP5;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OC NCBI_TaxID=9541;
OX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain parietal lobe;
RA Oseida N., Kusuda J., Hirata M., Tanuma R., Hida M., Sugano S.,
RA Hirai M., Hashimoto K.;
RT "Search for genes positively selected during primate evolution by 5'-
RT end-sequencing screening of cynomolgus monkey cDNAs.";
RL Genomics 79:657-662(2002).
DR EMBL; AB072025; BAB6814.1; -
DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.
DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.
DR GO; GO:0046961; F:hydrogen-transporting ATPase activity; roca. .; IEA.
DR GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.
DR InterPro: IPR006387; ATP synth_F6.
DR Pfam; PF05511; ATP-synt_F6; 1.
SQ SEQUENCE 108 AA; 12587 MW; A3AC44891716C1E CRC64;

Query Match 92.1%; Score 363; DB 2; Length 108;
Best Local Similarity 89.5%; Pred. No. 9e-29;
Matches 68; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 1 NKELDPIQKLFVDKIREYKSKRQTSGGPVDASSEYQDLERELFKLKMFGNADMTPT 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 33 NKELDPIQKLFVDKIREYKSKRQTSGGPVDGPEYQDLERELFKLKMFGKADMTPT 92
QY 61 FKPEDPKFEVLEKPOA 76
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 93 FKPEDPKFEVLEKPPA 108

RESULT 4
ATPR_BOVIN STANDARD; PRT; 108 AA.
ID ATPR_BOVIN
AC P02721;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE ATP synthase coupling factor 6, mitochondrial precursor (EC 3.6.3.14)
DE (F6).
GN Name=ATP5;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
[1]
RP SEQUENCE FROM N.A.
MEDLINE=88163536; PubMed=2694843;

```

```

RA Walker J.E., Gay N.J., Powell S.J., Kostina M., Dyer M.R.;
RT "ATP synthase from bovine mitochondria: sequences of imported
RT precursors of oligomycin sensitivity conferral protein, factor 6, and
RT adenosinetriphosphatase inhibitor protein.";
RL Biochemistry 26:8613-8619(1987).
[2]
RN SEQUENCE OF 33-108.
RX MEDLINE=85038563; PubMed=6149548;
RA Fang J.-K., Jacobs J.W., Kanner B.I., Rackner R., Bradshaw R.A.;
RT "Amino acid sequence of bovine heart coupling factor 6.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:6603-6607(1984).
[3]
RN SEQUENCE OF 33-37.
RP TISSUE=Heart;
RC MEDLINE=91242449; PubMed=1827992;
RA Walker J.E., Lutter R., Dupuis A., Runswick M.J.;
RT "Identification of the subunits of F1F0-ATPase from bovine heart
RT mitochondria.";
RL Biochemistry 30:5369-5378(1991).
CC -1- FUNCTION: This is one of the chains of the nonenzymatic component
CC (CF(0) subunit) of the mitochondrial ATPase complex. F6 seems to
CC be part of the stalk that links CF(0) to CF(1). Also involved in
CC the restoration of oligomycin-sensitive ATPase activity to
CC depleted F1-F0 complexes.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -1- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
CC core - and CF(0) - the membrane proton channel. CF(0) seems to
CC have nine subunits: a, b, c, d, e, f, g, F6 and 8 (or A6L).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M19217; AAA30511.1; -
DR PIR; B27382; ULB06.
DR InterPro; IPR008387; ATP_synth_F6.
DR Pfam; PF05511; ATP-synt_F6; 1.
DR CF(0); Direct protein sequencing; Hydrogen ion transport;
KW Mitochondrion; Transit peptide.
FT TRANSIT 1 32 Mitochondrion.
FT CHAIN 33 108 ATP synthase coupling factor 6.
FT CONFLICT 94 94 T -> F (in Ref. 2).
SQ SEQUENCE 108 AA; 12532 MW; E5376A0518C31C8 CRC64;

Query Match 87.8%; Score 346; DB 1; Length 108;
Best Local Similarity 82.9%; Pred. No. 4.5e-27;
Matches 63; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
QY 1 NKELDPIQKLFVDKIREYKSKRQTSGGPVDASSEYQDLERELFKLKMFGNADMTPT 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 33 NKELDPIQKLFVDKIREYKSKRQTSGGPVDGPEYQDLERELFKLKMFGKADMTPT 92
QY 61 FKPEDPKFEVLEKPOA 76
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 93 FTFEDPKFEVLEKPOA 108

RESULT 5
ATPR_PIG STANDARD; PRT; 76 AA.
ID ATPR_PIG
AC P1618;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE ATP synthase coupling factor 6, mitochondrial (EC 3.6.3.14) (F6).
GN Name=ATP5;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```





QY 1 NKELDPIQKLFVDKIREYKSKRTSGGVPDASSEYQOELERELPKLKOMFGNADMTPEPT 60  
 DB 33 NKELDPIQKLFVDKIREYKSKRTSGGVPDASSEYQOELERELPKLKOMFGNADMTPEPT 92  
 QY 61 FKEDPKFEVLEKPOA 76  
 DB 93 FKEDPKFEVLEKPOA 108

RESULT 7  
 ATRP RAT STANDARD; PRT; 108 AA.  
 ID ATRP RAT STANDARD; PRT; 108 AA.  
 AC P21571;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 05-JUN-2004 (Rel. 44, Last annotation update)  
 DE ATP synthase coupling factor 6, mitochondrial precursor (EC 3.6.3.14) (F6).  
 GN Name=Atp5f;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=91024964; PubMed=2145831;  
 RA Higuti T., Osaka F., Yoshihara Y., Tsurumi C., Kawamura Y., Tani I., Toda H., Kakuno T., Sakiyama F., Tanaka K., Ichihara A.;  
 RT "cDNA cloning and precursor of the import precursor of coupling factor 6 in H(+) -ATP synthase from rat liver mitochondria.";  
 RL Biochem. Biophys. Res. Commun. 171:1079-1086(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Anterior pituitary;  
 RX MEDLINE=92339904; PubMed=1386054; DOI=10.1016/0378-1119(92)90528-W;  
 RA Tracer H.L., Loh Y.P., Birch N.P.;  
 RT "rat mitochondrial coupling factor 6: molecular cloning of a cDNA encoding the imported precursor.";  
 RL Gene 116:291-292(1992).  
 RN [3]  
 RP SEQUENCE OF 33-66.  
 RC TISSUE=Liver;  
 RX MEDLINE=93054567; PubMed=1429613;  
 RA Higuti T., Yoshihara Y., Kuroiwa K., Kawamura Y., Toda H., Sakiyama F.;  
 RT "A simple, rapid method for purification of epsilon-subunit, coupling factor 6, subunit d, and subunit e from rat liver H(+)-ATP synthase and determination of the complete amino acid sequence of epsilon-subunit.";  
 RL J. Biol. Chem. 267:22658-22661(1992).  
 CC -1- FUNCTION: This is one of the chains of the nonenzymatic component (CF0) subunit of the mitochondrial ATPase complex. F6 seems to be part of the stalk that links CF0 to CF1.  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate + H(+) (out).  
 CC -1- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic core - and CF(0) - the membrane proton channel. CF(0) seems to have nine subunits: a, b, c, d, e, f, g, F6 and 8 (or A6L).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL: M73030; AAA40954.1; -;  
 CC EMBL: X54510; CAA36369.1; -;  
 CC PIR: JC1167; JC1167.  
 CC HSC-2DPAGE; P21571; RAT.

DR RGD; 621376; Atp5j.  
 DR InterPro: IPR008387; ATP\_synth\_F6.  
 DR Pfam: PF05511; ATP-synth\_F6; 1.  
 KM CF(0); Direct protein sequencing; Hydrogen ion transport;  
 KM Mitochondrion; Transit peptide.  
 FT TRANSIT 1 32  
 FT CHAIN 33 108  
 SQ SEQUENCE 108 AA; 12494 MW; PF1177C9681B5F51 CRC64;

Query Match 80.5%; Score 317; DB 1; Length 108;  
 Best Local Similarity 75.0%; Pred. No. 3.7e-24;  
 Matches 57; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

QY 1 NKELDPIQKLFVDKIREYKSKRTSGGVPDASSEYQOELERELPKLKOMFGNADMTPEPT 60  
 DB 33 NKELDPIQKLFVDKIREYKSKRTSGGVPDASSEYQOELERELPKLKOMFGNADMTPEPT 92  
 QY 61 FKEDPKFEVLEKPOA 76  
 DB 93 FKEDPKFEVLEKPOA 108

RESULT 8  
 ID 068FJ1 PRELIMINARY; PRT; 107 AA.  
 AC 068FJ1;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE MGC86324 protein.  
 GN Name=MGC86324;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W., Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altshuler S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Canninci P., Prange C., Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Guarnatone P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Vallation D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fehy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalek A.U., Smalhus D.E., Schnerch A., Schein J.E., Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Klein S., Gerhard D.S.;

RL Submitted (Aug-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC079787; AAH79787.1; -.  
DR InterPro: IPR008387; ATP synth\_F6.  
DR Pfam: PF05511; ATP-synt\_F6; 1.  
SQ SEQUENCE 107 AA; 12384 MW; C6EBBC7E08B854E CRC64;

Query Match 62.2%; Score 245; DB 2; Length 107;  
Best Local Similarity 61.8%; Pred. No. 6e-17;  
Matches 42; Conservative 14; Mismatches 12; Indels 0; Gaps 0;

QY 2 KEIDPIQKLFVDKIREYKSKRQTSGGPVDAASSEYQGLEBELFKLQKMFNADMTPTPT 61  
DB 36 KEIDPIQKLFVDKIREYNTKSKQKASGVDAAGSEYQKEMEDISKLRLYGGGDLTKRPPDF 95

QY 62 KFEDEPKFE 69  
DB 96 KFEDEPKFE 103

## RESULT 9

ID Q6PG55 PRELIMINARY; PRT; 107 AA.  
AC Q6PG55;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE MGC68738 protein.  
OS *Xenopus laevis* (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
OC Xenopodidae; *Xenopus*.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spine;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,  
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raba S.S., Loguelfano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RP SEQUENCE FROM N.A.  
RC TISSUE=Spine;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strauberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.;  
RT "Genetic and genomic tools for *Xenopus* research: The NIH *Xenopus*  
RT Initiative";  
RL Dev. Dyn. 225:384-391(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spine;  
RA Klein S., Strauberg R.;  
RL Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC057213; AAH57213.1; -.  
DR GO: GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.  
GO: GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.

DR GO: GO:0046961; F:hydrogen-transporting ATPase activity; rota. .; IEA.  
DR GO: GO:0015986; P:ATP synthesis coupled proton transport; IEA.  
DR InterPro: IPR008387; ATP synth\_F6.  
DR Pfam: PF05511; ATP-synt\_F6; 1.  
SQ SEQUENCE 107 AA; 12359 MW; 0F2C5BD80E6F286 CRC64;

Query Match 60.2%; Score 237; DB 2; Length 107;  
Best Local Similarity 60.3%; Pred. No. 3.8e-16;  
Matches 41; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 2 KEIDPIQKLFVDKIREYKSKRQTSGGPVDAASSEYQGLEBELFKLQKMFNADMTPTPT 61  
DB 36 KEIDPIQKLFVDKIREYNTKSKQKAGPVDAAGSEYQKEMEDISKLRLYGGGDLTKRPPDF 95

QY 62 KFEDEPKFE 69  
DB 96 KFEDEPKFE 103

## RESULT 10

ID Q6NF7 PRELIMINARY; PRT; 112 AA.  
AC Q6NF7;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE zgc:77541.  
GN ORFNames=zgc:77541.  
OS *Brachydanio rerio* (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,  
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raba S.S., Loguelfano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RA Strauberg R.;  
RL Submitted (Feb-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC066613; AAH66613.1; -.  
DR ZFIN; ZDB-GENE-040426-2534; zgc:77541.  
DR GO: GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.  
DR GO: GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.  
DR GO: GO:0046961; F:hydrogen-transporting ATPase activity; rota. .; IEA.  
DR GO: GO:0015986; P:ATP synthesis coupled proton transport; IEA.  
DR InterPro: IPR008387; ATP synth\_F6.  
DR Pfam: PF05511; ATP-synt\_F6; 1.  
SQ SEQUENCE 112 AA; 12244 MW; 4DA7B91BDAP87D7 CRC64;  
Query Match 46.2%; Score 182; DB 2; Length 112;

```
Best Local Similarity 53.2%; Pred. No. 1.3e-10;
Matches 33; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

OY 2 KELDPIQKLFVDKIREYKSKRQTSQGPVDASSEYQOELERELFKLKMFGNADMTTFPF 61
Db 37 KMDPIQKLFVDKIRYNSKSKVGVADGVPYKMLAEETTKLQRLYGGDLSPQPF 96

OY 62 KF 63
Db 97 SF 98

RESULT 11
O707P4 PRELIMINARY: PRT; 117 AA.
ID O707P4
AC O707P4
DT 01-MAR-2004 (TREMBLrel. 26, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DE AGCP4445 (Fragment).
GN Name=agCG56939; ORFNames=ENSNANG00000019097;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1 CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC
DR EMBL; AA01008952; EAA10598.1;
DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.
DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.
DR GO; GO:0046961; F:hydrogen-transporting ATP synthase activity; IEA.
DR GO; GO:0015966; P:ATP synthetase coupled proton transport; IEA.
DR InterPro: IPR008387; ATP synth_F6.
DR Pfam; PF05511; ATP-synt_F6; 1.
FT NON_TER
SQ SEQUENCE 117 AA; 12984 MW; 98DA5B85F9376E95 CRC64;

Query Match 45.7%; Score 180; DB 2; Length 117;
Best Local Similarity 51.4%; Pred. No. 2.2e-10;
Matches 37; Conservative 13; Mismatches 18; Indels 4; Gaps 2;

OY 2 KELDPIQKLFVDKIREYKSKRQTSQGPVDASSEYQOELERELFKLKMFGNADMTTF 57
Db 40 KASDPIQKLFVDKIREYNSKSKVGVADGVPYKMLAEETTKLQRLYGGDLSPQPF 96

OY 58 PPTFKEDPKFE 69
Db 100 PPAKFEEPKID 111

RESULT 12
ATPR_DROME STANDARD; PRT; 106 AA.
ID ATPR_DROME
AC Q24407; Q9VCNO;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE ATP synthase coupling factor 6, mitochondrial precursor (EC 3.6.3.14)
DE (F6).
GN Name=ATPsyn-F6; ORFNames=CG4442;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC TISSUE=Ovary;
RX MEDLINE=99168769; PubMed=10071211; DOI=10.1007/s004380050942;
RA Caggese C., Ragone G., Perrini B., Moschetti R., De Pinto V.,
RA Calzari R., Barsanti P.,
RT "Identification of nuclear genes encoding mitochondrial proteins:
RT isolation of a collection of D. melanogaster cDNAs homologous to
RT sequences in the Human Gene Index database."
RL Mol. Gen. Genet. 261:64-70(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoeking R.A., Galle R.F.,
RA Brandon R.C., Rogers Y.-H.C., Blazer J.G., Change M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieri S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibbegan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Paclet J.M.,
RA Palazzolo R., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sidon-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Splet B., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [3]
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Miara S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield B.U., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review."
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
CC -1- FUNCTION: This is one of the chains of the nonenzymatic component
CC (CF(0) subunit) of the mitochondrial ATPase complex. F6 seems to
CC be part of the stalk that links CF(0) to CF(1).
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -1- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
CC core - and CF(0) - the membrane proton channel. CF(0) seems to
CC have nine subunits: a, b, c, d, e, f, g, F6 and 8 (or A6L).
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X99665; CAA67979.1; -  
 DR EMBL; AE003743; AAF56127.1; -  
 DR InfAct; Q24407; -  
 DR FlyBase; FBgn0016119; ATPsyn-*cf6*.  
 DR InterPro; IPR008387; ATP\_synth\_F6.  
 DR Pfam; PF05511; ATP-synth\_F6; 1.  
 CF(0); Hydrogen ion transport; Mitochondrion; Transist peptide.  
 TRANSIT 1  
 FT CHAIN ? 106 ATP synthase coupling factor 6.  
 SQ SEQUENCE 106 AA; 11936 MW; 08E1B074EB34B94B CRC64;

Query Match 39.2%; Score 154.5; DB 1; Length 106;  
 Best Local Similarity 43.8%; Pred. No. 7e-08;  
 Matches 35; Conservative 19; Mismatches 19; Indels 7; Gaps 4;

QY 1 NKEHDPLOKLVDPKIREYKRSKRGSGP-VPASSEYQOELERELFKLKMPGN--ADMN 56  
 Db 29 NKASDPDLOQLFIDKREYKQK-SAGGKLVDSNPDIRKELTDLRVAKQFSGDKTDMT 86

QY 57 TPTFKFEDPKFE-VLEKPKQ 75  
 Db 87 KPFEPQFDPVQVDPITQAPQ 106

RESULT 13  
 Q9VZ72 PRELIMINARY; PRT; 147 AA.  
 AC Q9VZ72;  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)  
 DE CGJ2027-PA.  
 GN ORENAMES=CGJ2027;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers J.H., Blazey R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA Abell J.F., Agbayant A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,  
 RA Bokora D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport U.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Faraz C., Ferreira S., Fleischmann W.,  
 RA Foeller C., Gilbert A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jaiswal R., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshneft A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Rajatnikar K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodgett, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao C., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
 RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*  
 RT melanogaster euchromatic genome sequence.";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
 RT a genomic perspective.";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
 RT a genomic perspective.";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [10]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [11]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [14]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [15]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [16]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [17]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [18]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [19]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [20]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [21]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [22]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [23]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [24]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [25]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [26]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [27]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [28]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [29]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [30]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [31]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [32]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [33]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [34]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [35]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [36]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0

Qy 62 KFED 65  
 Db 85 KLPD 88

## RESULT 14

Q812D0 PRELIMINARY; PRT; 54 AA.  
 AC Q812D0;  
 DT 01-JUN-2003 (TREMBLrel. 24, Created)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE ATP synthase coupling factor VI (fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed:15199140; DOI=10.1128/MCB.24.13.5844-5849.2004;  
 RA Risteveki S., O'Leary D.A., Thornell A.P., Owen M.J., Kola I.,  
 RA Hertzog P.J.;  
 RT "The ETS Transcription Factor GABP(alpha) Is Essential for Early  
 RT Embryogenesis.";  
 RL Mol. Cell. Biol. 24:5844-5849(2004).  
 DR EMBL; AF346288; AAC27831.1; -.  
 DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.  
 DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.  
 DR GO; GO:0046961; F:hydrogen-transporting ATPase activity; IEA.  
 DR GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.  
 DR InterPro; IPR008387; ATP synth\_F6.  
 DR Pfam; PF05511; ATP-synt\_F6; 1.  
 DR NOW TER  
 FT SEQUENCE 54 AA; 6285 MW; 82BFF0707AA7A29C CRC64;  
 SQ

Query Match 28.2%; Score 111; DB 2; Length 54;  
 Best Local Similarity 95.5%; Pred. No. 0.00077;  
 Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NKELDPIQKLFVDKIREYKSR 22  
 Db 33 NKELDPVQKLFVDKIREYKSR 54

## RESULT 15

Q75UE1 PRELIMINARY; PRT; 1934 AA.  
 AC Q75UE1;  
 DT 05-JUL-2004 (TREMBLrel. 27, Created)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 DE Myosin heavy chain.  
 GN Name=LjMyHc1;  
 OS Lampetra japonica (Japanese lamprey) (Entosphenus japonicus).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;  
 OC Petromyzontiformes; Petromyzontidae; Leptocephala.  
 NCBI\_TaxID=94989;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kusakabe R., Takechi M., Tochinal S., Kuratani S.;  
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB126173; BABD1606.1; -.  
 DR HSSP; P24733; IKK7.  
 DR GO; GO:0016459; C:myosin; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0003774; F:motor activity; IEA.  
 DR InterPro; IPR000048; IQ region.  
 DR InterPro; IPR001609; Myosin\_head.  
 DR InterPro; IPR004009; Myosin\_N.  
 DR InterPro; IPR002928; Myosin\_tail.  
 DR Pfam; PF00063; Myosin\_head; 1.  
 DR Pfam; PF02736; Myosin\_N; 1.  
 DR Pfam; PF01576; Myosin\_tail\_1; 1.

DR PRINTS; PR00193; MYOSINHEAVY.  
 DR ProDom; PD000355; Myosin\_head; 1.  
 DR SMART; SM00242; MYSC; 1.  
 DR PROSITE; PS50096; IQ; 1.  
 SQ SEQUENCE 1934 AA; 222488 MW; 7D4261D4C35F85DD CRC64;

Query Match 20.2%; Score 79.5; DB 2; Length 1934;  
 Best Local Similarity 31.8%; Pred. No. 57;  
 Matches 21; Conservative 13; Mismatches 31; Indels 1; Gaps 1;

Qy 4 LDP1QKLFVDKIREYKSRQTSQGPVDAASSEYQOLEBELFKLKQMGADMTPTPKF 63  
 Db 1440 LDKKQKAFDKVLSWKQKFEESSQALEBAQKRSRLGTLPKLNAY-EETLDHLETPKR 1498

Qy 64 EDPKFE 69  
 Db 1499 ENKNLQ 1504

Search completed: April 4, 2005, 19:23:24  
 Job time : 120.5 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2005, 19:11:13 ; Search time 26 Seconds  
(without alignments)  
281.249 Million cell updates/sec

Title: US-09-831-951A-1

Perfect score: 394  
Sequence: 1 NKEIDPIQKLFVDKIREYKS.....TFPTFKPEDPKFEVLEKPKQA 76

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	392	99.5	108	2	UT0563
2	346	87.8	108	1	UJB06
3	342	86.8	76	2	S00212
4	317	80.5	108	2	JC1167
5	79.5	20.2	1940	1	S04090
6	78.5	19.9	557	2	A61256
7	77.5	19.7	764	2	I51302
8	76.5	19.4	1940	1	A24922
9	75.5	19.2	1931	2	A59234
10	75.5	19.2	1934	2	I48153
11	75.5	19.2	1935	1	A37102
12	75.5	19.2	1938	1	S06005
13	75.5	19.2	1938	2	I49464
14	75.5	19.2	1939	1	A46762
15	75.5	19.2	1939	2	I48175
16	75	19.0	15	2	PD0444
17	73.5	18.7	1935	1	S06006
18	72.5	18.4	285	2	J00448
19	72.5	18.4	741	2	S39082
20	72.5	18.4	936	2	S39083
21	72.5	18.4	955	2	S24348
22	72.5	18.4	1940	2	A29320
23	71.5	18.1	1039	2	S18199
24	71	18.0	555	2	T17320
25	71	18.0	1176	2	S40899
26	70.5	17.9	824	2	AB3354
27	70.5	17.9	1938	1	JX0178
28	69.5	17.6	541	2	AE0464
29	69.5	17.6	929	2	T28927

30	69.5	17.6	1937	2	I18055	myosin heavy chain
31	69	17.5	436	2	B87374	ReaA secretion sys
32	68.5	17.4	142	2	I50496	light meromyosin -
33	68	17.3	284	2	S17563	RNA-binding protei
34	68	17.3	342	2	A24263	myosin heavy chain
35	68	17.3	509	2	S45631	DNA primase chain
36	68	17.3	808	2	C72858	AcorF-66 protein -
37	68	17.3	1935	2	A59286	myosin heavy chain
38	67.5	17.1	302	2	S56751	single stranded D
39	67.5	17.1	353	1	S56750	single stranded D
40	67.5	17.1	2186	2	H89960	hypothetical prote
41	66.5	16.9	385	2	C24263	myosin heavy chain
42	65.5	16.6	793	1	SURFCA	endopeptidase Clp
43	65.5	16.6	1938	2	A59293	skeletal myosin he
44	65	16.5	143	2	A12057	hypothetical prote
45	65	16.5	217	2	G70407	hypothetical prote

## ALIGNMENTS

## RESULT 1

UT0563  
coupling factor 6 precursor, mitochondrial - human  
N/Alternate names: ATP synthase coupling factor 6  
C/Species: Homo sapiens (man)  
C/Date: 17-Apr-1993 #sequence revision 17-Apr-1993 #text\_change 09-Jul-2004  
C/Accession: UT0563; JQ1066; T14747  
R/Javed, A.A.; Ogata, K.; Sanadi, D.R.  
Gene 97, 307-310, 1991  
A/Title: Human mitochondrial ATP synthase: cloning cDNA for the nuclear-encoded precurs  
A/Reference number: UT0563; MUID:91153664; PMID:1825642  
A/Accession: UT0563

A/Molecule type: mRNA  
A/Residues: 1-108 <JAV>  
A/Cross-references: UNIPROT:P18859; GB:M37104; NID:g179274; PIDN:AAA51807.1; PID:g179277  
A/Experimental source: fetal muscle  
R/Higuti, T.; Tsunumi, C.; Kawamura, Y.; Tsujita, H.; Osaka, F.; Yoshihara, Y.; Tani, I  
Biochem. Biophys. Res. Commun. 178, 793-799, 1991  
A/Title: Molecular cloning of cDNA for the import precursor of human coupling factor 6  
A/Reference number: JQ1066; MUID:91315516; PMID:1830479  
A/Accession: JQ1066  
A/Molecule type: mRNA  
A/Residues: 1-67, 'H', 69-108 <HTG>  
A/Cross-references: GB:M73031; NID:g183785; PIDN:AAA58630.1; PID:g183786  
A/Experimental source: kidney  
R/Blum, H.; Baerach, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, August 1999  
A/Reference number: Z18179  
A/Accession: T14747  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-108 <BLU>  
A/Cross-references: EMBL:AL110183  
A/Experimental source: fetal kidney; clone DKFZ566A221  
C/Genetic8:  
A/Genes: GDB:ATP5J  
A/Cross-references: GDB:127519  
A/Map position: 2pter-2qter  
A/Note: DKFZ566A221.1  
C/Superfamily: coupling factor 6  
C/Keywords: mitochondrion; oxidative phosphorylation  
F1-32/Domain: transist peptide (mitochondrion) #status predicted <TPP>  
F133-108/Product: coupling factor 6 #status predicted <MAT>

Query Match 99.5%; Score 392; DB 2; Length 108;  
Best local similarity 98.7%; Pred. No. 4.9e-34;  
Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKEIDPIQKLFVDKIREYKSRQTSQGPVDASSSYQGLERELFKLQMGFNADMTPT 60  
DB 33 NKEIDPIQKLFVDKIREYKSRQTSQGPVDASSSYQGLERELFKLQMGFNADMTPT 92

OY 61 FKPEDPKFEVLEKPOA 76  
 |||||  
 Db 93 FKPEDPKFEVLEKPOA 108

## RESULT 2

coupling factor 6 precursor, mitochondrial - bovine  
 N:Alternate names: ATP synthase coupling factor 6  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 17-May-1995 #sequence\_revision 28-Oct-1994 #text\_change 09-Jul-2004  
 C/Accession: B27382; A03180; E39566  
 R/Walker, J.E.; Gay, N.J.; Powell, S.J.; Kostina, M.; Dyer, M.R.  
 Biochemistry 26, 8613-8619, 1987  
 A>Title: ATP synthase from bovine mitochondria: sequences of imported precursors of oligo  
 A/Reference number: A90527; MUID:88163536; PMID:2894843  
 A/Accession: B27382  
 A/Molecule type: mRNA  
 A/Residues: 1-108 <MUL>  
 A/Cross-references: UNIPROT:P02721; GB:M19217; NID:G163035; PIDN:AAA30511.1; PID:G163036  
 R/Pang, J.; Jacobs, J.W.; Kanner, B.I.; Racker, E.; Bradshaw, R.A.  
 Proc. Natl. Acad. Sci. U.S.A. 81, 6603-6607, 1984  
 A>Title: Amino acid sequence of bovine heart coupling factor 6.  
 A/Reference number: A03180; MUID:85038563; PMID:6149548  
 A/Accession: A03180  
 A/Molecule type: protein  
 A/Residues: 33-93; 'F', 95-108 <FAN>  
 A/Experimental source: heart  
 R/Walker, J.E.; Lutter, R.; Dupuis, A.; Runswick, M.J.  
 Biochemistry 30, 5369-5378, 1991  
 A>Title: Identification of the subunits of F-1F-0-ATPase from bovine heart mitochondria.  
 A/Reference number: A39566; MUID:91242449; PMID:1827992  
 A/Accession: E39566  
 A/Molecule type: protein  
 A/Residues: 33-37 <MA2>  
 C/Comment: This is one of the soluble components required for coupling of phosphorylation  
 C/Superfamily: coupling factor 6  
 C/Keywords: mitochondrion; oxidative phosphorylation  
 F:1-32/Domain: transmembrane peptide (mitochondrion) #status predicted <TNP>  
 F:33-108/Product: coupling factor 6 #status experimental <MAT>

Query Match 87.8%; Score 346; DB 1; Length 108;  
 Best Local Similarity 82.9%; Pred. No. 3.2e-29;  
 Matches 63; Conservative 8; Mismatches 5; Indels 0; Gaps 0;  
 OY 1 NKELDPIQKLFVFKIREYKSKROTSGGPDVASSGYOELRELFKIKOMFGNADMTPT 60  
 |||||  
 Db 33 NKELDPIQKLFVFKIREYKSKROTSGGPDVASSGYOELRELFKIKOMFGNADMTPT 92  
 |||||  
 OY 61 FKPEDPKFEVLEKPOA 76  
 |||||  
 Db 93 FKPEDPKFEVLEKPOA 108

## RESULT 3

S00212  
 coupling factor 6, mitochondrial - pig  
 C/Species: Sus scrofa domestica (domestic pig)  
 C>Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 09-Jul-2004  
 C/Accession: S00212  
 R/Chen, Z.W.; Mutt, V.; Barros-Soderling, J.; Joernvall, H.  
 FEBS Lett. 226, 43-46, 1987  
 A>Title: Isolation and structural characterization of porcine coupling factor 6 from int  
 A/Reference number: S00212; MUID:8803634; PMID:2961617  
 A/Accession: S00212  
 A/Molecule type: protein  
 A/Residues: 1-76 <CHE>  
 A/Cross-references: UNIPROT:P13618  
 C/Superfamily: coupling factor 6  
 C/Keywords: mitochondrion; oxidative phosphorylation

Query Match 86.8%; Score 342; DB 2; Length 76;  
 Best Local Similarity 81.6%; Pred. No. 5.6e-29;

Matches 62; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

OY 1 NKELDPIQKLFVFKIREYKSKROTSGGPDVASSGYOELRELFKIKOMFGNADMTPT 60  
 |||||  
 Db 1 NKELDPIQKLFVFKIREYKSKROTSGGPDVASSGYOELRELFKIKOMFGNADMTPT 92  
 |||||  
 OY 61 FKPEDPKFEVLEKPOA 76  
 |||||  
 Db 61 FKPEDPKFEVLEKPOA 108

## RESULT 4

coupling factor 6 precursor - rat  
 C/Species: Rattus norvegicus (Norway rat)  
 C>Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 09-Jul-2004  
 C/Accession: J01167; A35933; C44300; I52239; S23283  
 R/Tracer, H.L.; Loh, Y.P.; Birch, N.P.  
 Gene 116, 291-292, 1992  
 A>Title: Rat mitochondrial coupling factor 6: molecular cloning of a cDNA encoding the  
 A/Reference number: J01167; MUID:92339904; PMID:1386054  
 A/Accession: J01167  
 A/Molecule type: mRNA  
 A/Residues: 1-108 <TRA>  
 A/Cross-references: UNIPROT:P21571; EMBL:X54510; NID:G14214; PIDN:CAA38369.1; PID:G5999  
 R/Higuti, T.; Osaka, F.; Yoshihara, Y.; Tsunumi, C.; Kawamura, Y.; Tani, I.; Toda, H.;  
 Biochem. Biophys. Res. Commun. 171, 1079-1086, 1990  
 A>Title: cDNA cloning and sequencing for the import precursor of coupling factor 6 in H  
 A/Reference number: A35933; MUID:91024964; PMID:2145831  
 A/Accession: A35933  
 A/Molecule type: mRNA  
 A/Residues: 1-108 <HIG>  
 A/Cross-references: GB:M73030; NID:G203529; PIDN:AAA0954.1; PID:G203530  
 A/Note: part of this sequence, including the amino end of the mature protein, was deter  
 R/Higuti, T.; Yoshihara, Y.; Kuriwaka, K.; Kawamura, Y.; Toda, H.; Sakiyama, F.  
 J. Biol. Chem. 267, 22658-22661, 1992  
 A>Title: A simple, rapid method for purification of epsilon-subunit, coupling factor 6,  
 of epsilon-subunit.  
 A/Reference number: A44300; MUID:93054567; PMID:1429613  
 A/Accession: C44300  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 33-66 <H12>  
 A/Experimental source: liver  
 A/Note: sequence extracted from NCBI backbone (NCBI:119665)  
 C/Comment: This is one of the soluble components required for coupling of phosphorylation  
 C/Superfamily: coupling factor 6  
 C/Keywords: mitochondrion; oxidative phosphorylation  
 F:1-32/Domain: transmembrane peptide (mitochondrion) #status predicted <TNP>  
 F:33-108/Product: coupling factor 6 #status experimental <MAT>

Query Match 80.5%; Score 317; DB 2; Length 108;  
 Best Local Similarity 75.0%; Pred. No. 3.4e-26;  
 Matches 57; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

OY 1 NKELDPIQKLFVFKIREYKSKROTSGGPDVASSGYOELRELFKIKOMFGNADMTPT 60  
 |||||  
 Db 33 NKELDPIQKLFVFKIREYKSKROTSGGPDVASSGYOELRELFKIKOMFGNADMTPT 92  
 |||||  
 OY 61 FKPEDPKFEVLEKPOA 76  
 |||||  
 Db 93 FKPEDPKFEVLEKPOA 108

## RESULT 5

S04090  
 myosin heavy chain 3, skeletal muscle, embryonic - human  
 N:Contains: myosin ATPase (BC 3.6.4.1)  
 C/Species: Homo sapiens (man)  
 C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 C/Accession: S04090; S05446; S05442; S12460; S09333; A35082  
 R/Eiler, M.; Stedman, H.H.; Sylvestre, J.E.; Ferteles, S.H.; Rubinstein, N.A.; Kelly, A.A  
 Nucleic Acids Res. 17, 3591-3592, 1989



A>Title: Nucleotide sequence of full length human embryonic myosin heavy chain cDNA.  
 A/Reference number: S04090; MUID:89263803; PMID:2726495  
 A/Accession: S04090  
 A/Status: translation not shown  
 A/Molecule type: mRNA  
 A/Residues: 1-1940 <EEL>  
 A/Cross-references: UNIPROT:P11055; EMBL:X13988; NID:G34843; PIDN:CAA32167.1; PID:G34844  
 R/Editor: M.; Stedman, H.H.; Sylvester, J.E.; Fertels, S.H.; Wu, Q.L.; Raychowdhury, M.K.  
 FEBS Lett. 256, 21-28, 1989  
 A>Title: Human embryonic myosin heavy chain cDNA. Interspecies sequence conservation of  
 A/Reference number: S06146; MUID:9003298; PMID:2806546  
 A/Accession: S06146  
 A/Molecule type: mRNA  
 A/Residues: 774-1662, 'QT', 1665-1940 <EL2>  
 A/Cross-references: EMBL:X13100; NID:G31143; PIDN:CAA31492.1; PID:G31144  
 R/Karsch-Mizrachi, I.; Travis, M.; Blau, H.; Levinwand, L.A.  
 Nucleic Acids Res. 17, 6167-6179, 1989  
 A>Title: Expression and DNA sequence analysis of a human embryonic skeletal muscle myosin  
 A/Reference number: S05442; MUID:89366648; PMID:2771643  
 A/Accession: S05442  
 A/Molecule type: DNA  
 A/Residues: 856-1390, 'KK', 1393-1940 <KAR>  
 A/Cross-references: EMBL:X15696; NID:G36504; PIDN:CAA3731.1; PID:G133513  
 R/Stedman, H.H.; Ellier, M.; Jullian, E.H.; Fertels, S.H.; Sarkar, S.; Sylvester, J.E.; K  
 J. Biol. Chem. 265, 3568-3576, 1990  
 A>Title: The human embryonic myosin heavy chain. Complete primary structure reveals evol  
 A/Reference number: A35082; MUID:9015023; PMID:2303463  
 A/Accession: A35082  
 A/Contents: annotation; chromosomal assignment  
 R/Boder, E.  
 Submitted to the EMBL Data Library, January 1989  
 A/Reference number: S12458  
 A/Accession: S12458  
 A/Molecule type: mRNA  
 A/Residues: 856-1330, 'G', 1332-1390, 'KK', 1393-1607, 'RA', 1610-1940 <BOB>  
 A/Cross-references: EMBL:X15193; NID:G29463; PIDN:CAA35942.1; PID:G29464  
 A/Experimental source: clone gEMC-E  
 R/Boder, E.; Buchberger-Seidl, A.; Braun, T.; Singh, S.; Goedde, H.W.; Arnold, H.H.  
 Eur. J. Biochem. 189, 55-65, 1990  
 A>Title: Identification of three developmentally controlled isoforms of human myosin hea  
 A/Reference number: S09331; MUID:90235862; PMID:1691980  
 A/Accession: S09331  
 A/Molecule type: mRNA  
 A/Residues: 856-901, 'X', 903-971, 'X', 973-1041, 'X', 1043-1111, 'X', 1113-1181, 'X', 1183-1251, '  
 1, 'X', 1673-1741, 'X', 1743-1811, 'X', 1813-1881, 'X', 1883-1940 <BOB>  
 A/Cross-references: EMBL:X15193  
 A/Accession: X15193  
 A/Contents: GDB:MYH3  
 A/Cross-references: GDB:119443; OMIM:160720  
 A/Map position: 17p13.1-17p13.1  
 C/Superfamily: myosin heavy chain; myosin motor domain homology  
 C/Keywords: actin binding; ATP; coiled coil; hydrolyase; methylated amino acid; muscle co  
 F/89-767/Domain: myosin motor domain homology <MOT>  
 F/119-186/Region: nucleotide-binding motif A (P-loop)  
 F/519-586/Region: actin binding #status predicted  
 F/656-678/Region: actin binding #status predicted  
 F/840-1940/Domain: coiled coil #status predicted <COI>  
 F/840-1280/Region: 52  
 F/181-1940/Region: light meromyosin  
 F/130/Modified site: N6,N6,N6-trimethyllysine (lys) #status predicted  
 F/185/Binding site: ATP (lys) #status predicted  
 F/696,706/Active site: Cys #status predicted

Query Match 20.2%; Score 79.5; DB 1; Length 1940;  
 Best Local Similarity 33.3%; Pred. No. 6.1;  
 Matches 22; Conservative 13; Mismatches 30; Indels 1; Gaps 1;

4 LDPIQKLVNDKIRKYSKRTSGGPVDSSEYQOELRELFKTKOMFGNADMTPTPKF 63  
 Db 1443 LDKQKRFDPKVLAWKTKCESQAELESLKESRSLSTELFKLKNAYBEA-LDQLETVKR 1501

QY 64 EDPKFE 69  
 Db 1502 ENKYLE 1507

RESULT 6  
 A61256  
 myosin heavy chain, skeletal muscle - rabbit (fragment)  
 C/Species: Oryctolagus cuniculus (domestic rabbit)  
 C/Date: 12-May-1994 #sequence\_revision 12-May-1994 #text\_change 09-Jul-2004  
 A/Accession: A61256  
 R/Jackson, S.J.; Stewart, M.  
 J. Cell Sci. 99, 823-836, 1991  
 A>Title: Expression in Escherichia coli of fragments of the coiled-coil rod domain of r  
 A/Reference number: A61256; MUID:92121252; PMID:1770009  
 A/Accession: A61256  
 A/Molecule type: preliminary  
 A/Status: preliminary  
 A/Residues: 1-557 <ATK>  
 A/Cross-references: UNIPROT:Q7M3B1; GB:X59602  
 C/Superfamily: myosin heavy chain; myosin motor domain homology  
 C/Keywords: ATP; skeletal muscle

Query Match 19.9%; Score 78.5; DB 2; Length 557;  
 Best Local Similarity 33.3%; Pred. No. 1.9;  
 Matches 22; Conservative 13; Mismatches 30; Indels 1; Gaps 1;

4 LDPIQKLVNDKIRKYSKRTSGGPVDSSEYQOELRELFKTKOMFGNADMTPTPKF 63  
 Db 60 LDKQKRFDPKVLAWKTKCESQAELESLKESRSLSTELFKLKNAYBEA-LDQLETVKR 118

QY 64 EDPKFE 69  
 Db 119 ENKYLE 124

RESULT 7  
 I51302  
 myosin heavy chain - chicken (fragment)  
 C/Species: Gallus gallus (chicken)  
 C/Date: 04-Sep-1997 #sequence\_revision 04-Sep-1997 #text\_change 09-Jul-2004  
 A/Accession: I51302  
 R/Rutzy, K.E.; Rhee, J.T.; Bader, D.  
 Development 120, 871-883, 1994  
 A>Title: Expression of the atrial-specific myosin heavy chain AMHC1 and the establishme  
 A/Reference number: I51302; MUID:95324374; PMID:7600964  
 A/Accession: I51302  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-764 <YDT>  
 A/Cross-references: UNIPROT:Q91411; GB:S78540; NID:G1000404; PIDN:AA834772.1; PID:G1000  
 C/Superfamily: myosin heavy chain; myosin motor domain homology

Query Match 19.7%; Score 77.5; DB 2; Length 764;  
 Best Local Similarity 33.3%; Pred. No. 3.4;  
 Matches 22; Conservative 12; Mismatches 31; Indels 1; Gaps 1;

4 LDPIQKLVNDKIRKYSKRTSGGPVDSSEYQOELRELFKTKOMFGNADMTPTPKF 63  
 Db 271 LDKQKRFDPKVLAWKTKCESQAELESLKESRSLSTELFKLKNAYBEA-LDQLETVKR 329

QY 64 EDPKFE 69  
 Db 330 ENKYLE 335

RESULT 8  
 A24922  
 myosin heavy chain, skeletal muscle, embryonic - rat  
 N/Contains: myosin ATPase (BC 3.6.4.1)  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 A/Accession: A24922; B24263  
 R/Strehler, E.E.; Strehler-Pag, M.A.; Perriard, J.C.; Periaamy, M.; Nadal-Ginard, B.  
 J. Mol. Biol. 190, 291-317, 1986  
 A>Title: Complete nucleotide and encoded amino acid sequence of a mammalian myosin heavy

A:Reference number: A24922; MUID:87060988; PMID:3783701  
A:Accession: A24922  
A:Molecule type: DNA  
A:Residues: 1-1940 <STR>  
C:Cross-references: UNIPROT:P12847; GB:X04267; GB:X05004; NID:g56658; PIDN:CMA27817.1; E:R.Strehler, E.E.; Mahavi, V.; Perlasamy, M.; Nadal-Ginard, B.  
J. Biol. Chem. 260, 468-471, 1985  
A>Title: Intron positions are conserved in the 5' end region of myosin heavy-chain genes  
A:Reference number: A22538; MUID:85080119; PMID:2981212  
A:Accession: A22538  
A:Molecule type: DNA  
A:Residues: 1-168 <ST2>  
C:Cross-references: GB:L00370; GB:M10135; NID:g205860; PIDN:AAA41655.1; PID:g554476  
R:Perlasamy, M.; Wdyto, R.W.; Strehler-Page, M.A.; Strehler, E.E.; Nadal-Ginard, B.  
J. Biol. Chem. 260, 15856-15862, 1985  
A>Title: Characterization of cDNA and genomic sequences corresponding to an embryonic myosin motor domain homologous to chicken skeletal muscle myosin II  
A:Reference number: A24263; MUID:86059474; PMID:2999140  
A:Accession: B24263  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1358-1490, 'G' <PER>  
C:Cross-references: GB:K03468; NID:g205573; PIDN:AAA41652.1; PID:g205574  
A:Experimental source: clone pMHC-72  
C:Genetics:  
A:Introns: 68/3; 116/3; 169/1  
A>Note: The list of intron positions may be incomplete  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; muscle cd  
F:89-767/Domains: myosin motor domain homology <MWOT>  
F:179-166/Region: nucleotide-binding motif A (P-loop)  
F:549-566/Region: actin binding #status predicted  
F:685-678/Region: actin binding #status predicted  
F:840-1940/Domains: coiled coil #status predicted <COI>  
F:1281-1940/Region: light meromyosin  
F:1303/Modified site: Ng,N6,N6-trimethyllysine (lys) #status predicted  
F:188/Binding site: ATP (lys) #status predicted  
F:696,706/Active site: Cys #status predicted

Query Match 19.4%; Score 76.5; DB 1; Length 1940;  
Best Local Similarity 31.8%; Pred. No. 13;  
Matches 21; Conservative 14; Mismatches 30; Indels 1; Gaps 1;

Oy 4 LDPIQLFVDIKIREYKSQRQTSGGVPDASSEYQGELRELIFKLKOMFGNDMTFPFKF 63  
Db 1443 LDKQRNFDDKIATWKTCESSOAELEALKESSLSTEFLKNAYEEA-LDGLTVKR 1501  
| | : | : | : | : | : | : | : | : | : | : | : | : |  
Oy 64 EDPKFE 69  
Db 1502 ENKNLE 1507  
| | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 9  
A59234  
S:low myosin heavy chain 3 - quat1  
C:Species: Columba coturnix  
C:Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 08-Sep-2000  
C:Accession: A59234  
R:Nikovits Jr., W.; Wang, G.F.; Feldman, J.L.; Miller, J.B.; Wade, R.; Nelson, L.; Stock  
J. Biol. Chem. 271, 17047-17056, 1996  
A>Title: Isolation and characterization of an avian slow myosin heavy chain gene expressed  
A:Reference number: A59234; MUID:96291845; PMID:8663323  
A:Accession: A59234  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-1931 <NIK>  
C:Cross-references: GB:H53862; NID:g1289513; PIDN:AAC59912.1; PID:g1289514  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
F:81-761/Domains: myosin motor domain homology <MMO>

Query Match 19.2%; Score 75.5; DB 2; Length 1931;  
Best Local Similarity 31.8%; Pred. No. 16;  
Matches 21; Conservative 12; Mismatches 32; Indels 1; Gaps 1;

```

QY      4 LDPIQKLFVDKIREKSKSRQTSGGPVDASSSEYQGLRELEFKLKMPGNADMTPTPEKF 63
DB      1437 LDKKQNPDKITLSEMKQKFEESQTELEASQKFAKSLSTLEFKLKNAV - EESLEHLETFKR 1495
QY      64 EDPKFE 69
DB      1496 EKKNLQ 1501

RESULT 10
148153 myosin heavy chain beta, cardiac muscle [similarity] - golden hamster
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C/Accession: 148153; A28298
R/Wang, R./ Sole, M.J./ Culikerman, E./ Liew, C.C.
J. Mol. Cell. Cardiol. 26, 1155-1165, 1994
A/Title: Characterization and nucleotide sequence of the cardiac alpha-myosin heavy cha
A/Reference number: 148153; MUID:95115033; PMID:7815459
A/Accession: 148153
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1934 <RMS>
A/Cross-references: UNIPROT:P13540; GB:IL12104; NID:g402371; PIDN:AA62313.1; PID:g40237
R.Jandarek, M.A./ Sole, M.J./ Liew, C.C.
Nucleic Acids Res. 16, 4737, 1988
A/Title: Sequence of cDNA encoding the Syrian hamster cardiac beta-myosin heavy chain.
A/Reference number: A28298; MUID:88247788; PMID:3380703
A/Accession: A28298
A/Molecule type: mRNA
A/Residues: 962-965, 'E', 967-980, 'E', 981-985, 'Q', 987-1007, 'A', 1009, 'E', 1011, 'RKT', 1015-1
536, 'L', 1538-1555, 'K', 1557-1934 <JMN>
A/Cross-references: GB:X07273; NID:g49640; PIDN:CA30256.1; PID:g49641
A/Note: The authors translated the codon GTG for residue 1504 as Leu
C/Genetics:
A/Intons: 66/3; 114/3; 167/1; 176/2; 212/3; 243/3; 265/1; 298/1; 332/3; 379/1; 418/3;
23/3; 1389/2; 1450/3; 1506/1; 1547/3; 1650/3; 1718/3; 1760/3; 1852/3; 1884/3; 1929/3
C/Superfamily: myosin heavy chain; myosin motor domain homology
C/Keywords: actin binding; ATP; cardiac muscle; coiled coil; heart; muscle; nucleotide
F/87-184/Region: myosin motor domain homology <MOTOR>
F/87-184/Region: nucleotide-binding motif A (P-loop)

Query Match      19.2%; Score 75.5; DB 2; Length 1934;
Best Local Similarity 30.3%; Pred. No. 16;
Matches 20; Conservative 13; Mismatches 32; Indels 1; Gaps 1;

QY      4 LDPIQKLFVDKIREKSKSRQTSGGPVDASSSEYQGLRELEFKLKMPGNADMTPTPEKF 63
DB      1441 LDKKQNPDKITLSEMKQKFEESQTELEASQKFAKSLSTLEFKLKNAV - EESLEHLETFKR 1499
QY      64 EDPKFE 69
DB      1500 EKKNLQ 1505

RESULT 11
A37102
myosin beta heavy chain, cardiac and skeletal muscle - human
N/Contains: myosin ATPase (EC 3.6.4.1)
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C/Accession: A37102; S12733; J34224; B28908; A29497; A27558; I54254; S12458; S09331; SO.
R.Jandarek, T./ Diederich, K.W./ Haas, W./ Schleicher, U./ Lichte, F./ Pfordt, M./ Bach,
Genomics 8, 194-206, 1990
A/Title: The complete sequence of the human beta-myosin heavy chain gene and a comparat
A/Reference number: A37102; MUID:91065634; PMID:2249844
A/Accession: A37102
A/Molecule type: DNA
A/Residues: 1-1935 <JAB>
A/Cross-references: UNIPROT:P12883; GB:M57965; GB:M30603; NID:g179507; PIDN:AA51837.1;
R.Jewell, C.C./ Sole, M.J./ Yamauchi-Takahara, K./ Kellam, B./ Anderson, D.H./ Lin, L./ L
Nucleic Acids Res. 18, 3647-3651, 1990

```

A:Title: Complete sequence and organization of the human cardiac beta-myosin heavy chain  
 A:Reference number: S12733; MUID:90301496; PMID:2362820  
 A:Accession: S12733  
 A:Molecule type: DNA  
 A:Residues: 1-106, 'E', 108-671, 'LYH', 675-857, 'A', 859-941, 'NV', 944-1123, 'A', 1125-1158, 'C',  
 A:Cross-references: EMBL:X52889; NID:929726; PIDN:CAA37068.1; PID:929727  
 R:Yamauchi-Takahara, K.; Sole, M.J.; Llew, J.; Ing, D.; Llew, C.C.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 3504-3508, 1989  
 A:Title: Characterization of human cardiac myosin heavy chain genes.  
 A:Reference number: A94224; MUID:89264452; PMID:2726733  
 A:Accession: A94224  
 A:Molecule type: DNA  
 A:Residues: 1-87, 'Q', 89-106, 'E', 108-177, 1325-1702, 'DR', 1705-1786, 1788-1803, 'E', 1804-1935  
 R:Yamauchi-Takahara, K.; Sole, M.J.; Llew, J.; Ing, D.; Llew, C.C.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 7416-7417, 1989  
 A:Reference number: A94226  
 A:Contents: annotation; erratum  
 R:Kunadyabhi, M.; Tsuchimochi, H.; Komuro, I.; Takaku, F.; Yazaki, Y.  
 J. Clin. Invest. 82, 524-531, 1988  
 A:Title: Molecular cloning and characterization of human cardiac alpha- and beta-form my  
 human atrium.  
 A:Reference number: A92770; MUID:88299163; PMID:2965919  
 A:Accession: B28908  
 A:Molecule type: mRNA  
 A:Residues: 1412-1518, 'R', 1520-1574, 'NV', 1577-1935 <KOR>  
 A:Cross-references: GB:M21655  
 A:Note: the authors translated the codon AGC for residue 108 as Arg  
 R:Lichter, P.; Umeda, P.K.; Levin, J.E.; Vosberg, H.P.  
 Eur. J. Biochem. 160, 419-426, 1986  
 A:Title: Partial characterization of the human beta-myosin heavy-chain gene which is exp  
 A:Reference number: A24997; MUID:87030293; PMID:3021460  
 A:Accession: A24997  
 A:Molecule type: DNA  
 A:Residues: 682-721, 975-1112, 1854-1935 <LIC>  
 A:Cross-references: GB:X04627  
 R:Saez, L.; Gianola, K.M.; McNally, E.M.; Feghali, R.; Eddy, R.; Shows, T.B.; Leinwand  
 Nucleic Acids Res. 15, 5443-5459, 1987  
 A:Title: Human cardiac myosin heavy chain genes and their linkage in the genome.  
 A:Reference number: A93669; MUID:87260010; PMID:3037493  
 A:Accession: A27858  
 A:Molecule type: DNA  
 A:Residues: 1854-1865, 'A', 1867-1935 <SAB>  
 A:Cross-references: GB:X05631; GB:Y00362; NID:934643; PIDN:CAA29119.1; PID:934644  
 R:Dieterich, K.W.; Eiselle, I.; Ried, T.; Jaenicke, T.; Lichter, P.; Vosberg, H.P.  
 Hum. Genet. 81, 214-220, 1989  
 A:Title: Isolation and characterization of the complete human beta-myosin heavy chain ge  
 A:Reference number: 154254; MUID:89154425; PMID:2522082  
 A:Accession: 154254  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 653-720 <RES>  
 A:Cross-references: GB:M27636; NID:9179511; PIDN:AAA79019.1; PID:9601916  
 R:Bober, E.  
 submitted to the EMBL Data Library, January 1989  
 A:Reference number: S12458  
 A:Accession: S12458  
 A:Molecule type: mRNA  
 A:Residues: 785-1076, 'E', 1078-1123, 'A', 1125-1702, 'DE', 1705-1935 <BOB>  
 A:Cross-references: EMBL:X51591; NID:929467; PIDN:CAA35940.1; PID:929468  
 R:Bober, E.; Buchberger-Seidl, A.; Braun, T.; Singh, S.; Goede, H.W.; Arnold, H.H.  
 Eur. J. Biochem. 189, 55-65, 1990  
 A:Title: Identification of three developmentally controlled isoforms of human myosin hea  
 A:Reference number: S09331; MUID:90235862; PMID:1691980  
 A:Accession: S09331  
 A:Molecule type: mRNA  
 A:Residues: 785-830, 'X', 832-900, 'X', 902-970, 'X', 972-1040, 'X', 1042-1076, 'E', 1078-1110, 'X',  
 '1602-1670, 'X', 1672-1702, 'DE', 1705-1740, 'X', 1742-1810, 'X', 1812-1935 <BOB>  
 A:Cross-references: EMBL:X51591  
 R:Jandreski, M.A.; Llew, C.C.  
 Hum. Genet. 76, 47-53, 1987  
 A:Title: Construction of a human ventricular cDNA library and characterization of a beta  
 A:Reference number: S02229; MUID:87192738; PMID:3032769  
 A:Accession: S02229

A:Molecule type: mRNA  
 A:Residues: 1393-1702, 'DR', 1705-1935 <JAN>  
 A:Cross-references: EMBL:X06976; NID:934860; PIDN:CAA30039.1; PID:9825694  
 R:Saez, L.; Leinwand, L.A.  
 Nucleic Acids Res. 14, 2951-2969, 1986  
 A:Title: Characterization of diverse forms of myosin heavy chain expressed in adult hu  
 A:Reference number: A93616; MUID:86176778; PMID:2421254  
 A:Accession: B23767  
 A:Molecule type: mRNA  
 A:Residues: 'LUGVEELASG', 1311-1312, 'G', 1314-1355, 'R', 1357-1358, 'GD', 1361-1438, 'LQ', 1443  
 A:Note: the first ten codons of the sequence figure show the reverse complementary stre  
 C:Genetics:  
 A:Gene: GDB:MYH7  
 A:Cross-references: GDB:120215; OMIM:160760  
 A:Map position: 14q12-14q12  
 A:Introns: 67/3, 115/3, 168/1, 177/2, 213/3, 244/3, 266/1, 299/1, 333/3, 380/1, 419/3,  
 243/3, 1390/2, 1451/3, 1507/1, 1548/3, 1651/3, 1719/3, 1761/3, 1853/3, 1930/3  
 C:Superfamily: myosin heavy chain; myosin motor domain homology  
 C:Keywords: actin binding; ATP; cardiac muscle; coiled coil; heart; hydrolase; methylat  
 F:88-766/Domain: myosin motor domain homology <MMOT>  
 F:178-185/Region: nucleotide-binding motif A (P-loop)  
 F:548-585/Region: actin binding #status predicted  
 F:655-677/Region: actin binding #status predicted  
 F:839-1935/Domain: coiled coil #status predicted <COI>  
 F:839-1279/Region: S2  
 F:1280-1935/Region: light meromyosin  
 F:128/Modified site: N6,N6,N6-trimethyllysine (lys) #status predicted  
 F:184/Binding site: ATP (lys) #status predicted  
 F:695,705/Active site: Cys #status predicted  
 Query Match 19.2%; Score 75.5; DB 1; Length 1935;  
 Best Local Similarity 30.3%; Pred. No. 16;  
 Matches 20; Conservative 13; Mismatches 32; Indels 1; Gaps 1;  
 Oy 4 LDPIQKLVFDKIREYKSKRQTSQGPVADSEYQELRELFKLMQMGNDMTPTPKF 63  
 Db 1442 LDKQRNFDKILAEKQYKESQSELSQSEARSLSTELFKLNAV EESLEHLETFKR 1500  
 Oy 64 EDPKFE 69  
 Db 1501 ENKNLQ 1506  
 RESULT 12  
 S06005  
 myosin alpha heavy chain, cardiac muscle [similarity] - rat  
 N:Alternate names: alpha-myosin heavy chain  
 N:Contains: myosin ATPase (EC 3.6.4.1)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 C:Accession: S06005; S07535; A20971; A02988; I53305  
 R:McNally, E.M.; Gianola, K.M.; Leinwand, L.A.  
 Nucleic Acids Res. 17, 7527-7528, 1989  
 A:Title: Complete nucleotide sequence of full length cDNA for rat alpha cardiac myosin  
 A:Reference number: S06005; MUID:90016822; PMID:2758111  
 A:Accession: S06005  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-1938 <MCN>  
 A:Cross-references: UNIPROT:P02563; EMBL:X15938; NID:956654; PIDN:CAA34064.1; PID:95665  
 R:McNally, E.M.; Kraft, R.; Bravo-Zehnder, M.; Taylor, D.A.; Leinwand, L.A.  
 J. Mol. Biol. 210, 665-671, 1989  
 A:Title: Full-length rat alpha and beta cardiac myosin heavy chain sequences. Compariso  
 A:Reference number: S07535; MUID:90133919; PMID:2614840  
 A:Accession: S07535  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-1938 <MC2>  
 R:Mandavi, V.; Chambers, A.P.; Nadal-Ginard, B.  
 Proc. Natl. Acad. Sci. U.S.A. 81, 2626-2630, 1984  
 A:Title: Cardiac alpha- and beta-myosin heavy chain genes are organized in tandem.  
 A:Reference number: A20971; MUID:84194059; PMID:6585819  
 A:Accession: A20971

A:Molecule type: protein  
 A:Residues: 112,'A',14-45,'A',47-50,'AP',53-61,'E',83-86,'Q',88-109,111-133,'H',135-16  
 R,Mahdavi, V.; Periasamy, M.; Nadal-Ginard, B.  
 Nature 297, 659-664, 1982  
 A:Title: Molecular characterization of two myosin heavy chain genes expressed in the adu  
 A:Reference number: A02988; MUID:82220036; PMID:7045682  
 A:Accession: A02988  
 A:Molecule type: mRNA  
 A:Residues: 1512-1574,'S',1576-1721,'N',1723-1851,'N',1853-1869,'N',1871-1933,'I',1935-1  
 A>Note: there are 10 or more myosin heavy chain genes in the rat, at least two of which  
 R,Mahdavi, V.; Lompre, A.M.; Chambers, A.P.; Nadal-Ginard, B.  
 Eur. Heart J. 5, 181-191, 1984  
 A:Title: Cardiac myosin heavy chain isozyemic transitions during development and under pa  
 A:Reference number: I53305; MUID:85179510; PMID:6241892  
 A:Accession: I53305  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1872-1933,'I',1935-1938 <RES>  
 A:Cross-references: GB:M32697; NID:g205596; PIDN:AAA1658.1; PID:g205597  
 C:Superfamily: myosin heavy chain; myosin motor domain homology  
 C:Keywords: actin binding; ATP; cardiac muscle; coiled coil; heart; hydrolase; methylate  
 F:87-767/Domain: myosin motor domain homology <MMOT>  
 F:177-184/Region: nucleotide-binding motif A (P-loop)  
 F:548-585/Region: actin binding #status predicted  
 F:656-678/Region: actin binding #status predicted  
 F:840-1938/Domain: coiled coil #status predicted <COI>  
 F:840-1280/Region: S2  
 F:1281-1938/Region: light meromyosin  
 F:128/Modified site: N6,N6,N6-trimethyllysine (Lys) #status predicted  
 F:183/Binding site: ATP (Lys) #status predicted  
 F:686,706/Active site: Cys #status predicted

Query Match 19.2%; Score 75.5; DB 1; Length 1938;  
 Best Local Similarity 30.3%; Pred. No. 16;  
 Matches 20; Conservative 13; Mismatches 32; Indels 1; Gaps 1;

QY 4 LDPDQKLPVDFKIRYKSKQTSQSGPVDASSEYQELERELFKQKMGFNADMTTFPFK 63  
 DB 1443 LDKKQRFDFKILAEWKQKTESQSESSQKARSLSLELFKKNAY-EESLEHLETFK 1501

QY 64 EDPKFE 69  
 DB 1502 ENKMLQ 1507

# RESULT 13

149464

A:Molecule type: protein  
 C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
 C:Accession: I49464; I49463; I49462; I49461; I49604

R:Quinn-Laquer, B.K.; Kennedy, J.E.; Wei, S.C.; Beisel, K.W.  
 Genomics 13, 176-188, 1992

A:Title: Characterization of the allelic differences in the mouse cardiac alpha-myosin H  
 A:Reference number: A38207; MUID:92250040; PMID:1577481

A:Accession: I49464  
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA  
 A:Residues: 1-1938 <RES>

A:Cross-references: UNIPROT:Q02566; GB:M76601; NID:g191623; PIDN:AAA37162.1; PID:g191624

A:Accession: I49463  
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA  
 A:Residues: 1-193,'D',195-837,'S',839-955,'N',957-1938 <RES2>

A:Cross-references: GB:M76600; NID:g191621; PIDN:AAA37161.1; PID:g191622

A:Accession: I49462  
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA  
 A:Residues: 1-1938 <RES3>

A:Cross-references: GB:M76599; NID:g191619; PIDN:AAA37160.1; PID:g191620

A:Accession: I49461  
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-544,'A',546-1938 <RE4>  
 A:Cross-references: GB:M76598; NID:g191617; PIDN:AAA37159.1; PID:g191618  
 R,Glick, J.; Subramaniam, A.; Neumann, J.; Robbins, J.  
 U. Biol. Chem. 266, 9180-9185, 1991  
 A:Title: Isolation and characterization of the mouse cardiac myosin heavy chain genes.  
 A:Reference number: I49604; MUID:91225025; PMID:2026617  
 A:Accession: I49604  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-67 <RES5>  
 A:Cross-references: GB:M62404; NID:g192609; PIDN:AAA37424.1; PID:g192610  
 C:Superfamily: myosin heavy chain; myosin motor domain homology  
 C:Keywords: ATP; cardiac muscle; heart; nucleotide binding; P-loop  
 F:88-768/Domain: myosin motor domain homology <MMOT>  
 F:178-185/Region: nucleotide-binding motif A (P-loop)

Query Match 19.2%; Score 75.5; DB 2; Length 1938;  
 Best Local Similarity 30.3%; Pred. No. 16;  
 Matches 20; Conservative 13; Mismatches 32; Indels 1; Gaps 1;

QY 4 LDPDQKLPVDFKIRYKSKQTSQSGPVDASSEYQELERELFKQKMGFNADMTTFPFK 63  
 DB 1444 LDKKQRFDFKILAEWKQKTESQSESSQKARSLSLELFKKNAY-EESLEHLETFK 1502

QY 64 EDPKFE 69  
 DB 1503 ENKMLQ 1508

# RESULT 14

A46762  
 myosin alpha heavy chain, cardiac muscle - human

N:Contains: myosin ATPase (EC 3.6.4.1)

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 C:Accession: A46762; B46762; A49354; S18830; B32562; B33835; B27858; A28908

R:Matsumoto, R.; Beisel, K.W.; Furutani, M.; Arai, S.; Takao, A.  
 Am. J. Med. Genet. 41, 537-547, 1991

A:Title: Complete sequence of human cardiac alpha-myosin heavy chain gene and amino ac  
 A:Reference number: A46762; MUID:92133665; PMID:1776652

A:Accession: A46762  
 A:Status: nucleic acid sequence not shown

A:Molecule type: DNA  
 A:Residues: 1-1939 <MAT>

A:Cross-references: UNIPROT:P13533; UNIPROT:Q9UQV; DDBJ:D00943; NID:g219523; PIDN:BA01

A:Accession: B46762  
 A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA  
 A:Residues: 1-1461 <MA2>

R:BP, T.A.; Dixon, I.M.C.; Wang, H.Y.; Sole, M.J.; Liew, C.C.  
 Genomics 18, 505-509, 1993

A:Title: Structural organization of the human cardiac alpha-myosin heavy chain gene (MT  
 A:Reference number: A49354; MUID:9410346; PMID:8307559

A:Accession: A49354  
 A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA  
 A:Residues: 1-87,'Q',89-573,'Q',575-607,'A',609-743,'T',745-789,'W',791-1013,'V',1015-11

-1939 <BP>  
 A:Cross-references: GB:Z20656; NID:g297023; PIDN:CAA79675.1; PID:g297024

R:Brand, N.J.; Dabhadre, N.; Yacoub, M.; Barton, P.J.R.  
 Biochem. Biophys. Res. Commun. 179, 1255-1258, 1991

A:Title: Determination of the 5' exon structure of the human cardiac alpha-myosin heavy  
 A:Reference number: S18830; MUID:92028855; PMID:1930170

A:Accession: S18830  
 A:Status: preliminary

A:Molecule type: mRNA  
 A:Residues: 1-32 <BR>

A:Cross-references: EMBL:X56181; NID:g28318; PIDN:CAA39642.1; PID:g28319

R,Yamauchi-Takahara, K.; Sole, M.J.; Liew, J.; Ing, D.; Liew, C.C.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 3504-3508, 1989

A:Title: Characterization of human cardiac myosin heavy chain genes.  
 A:Reference number: A94224; MUID:89264452; PMID:2726733

A:Accession: B32562

A: Molecule type: DNA  
A: Residues: 1-87, 'Q', 89-177; 1551-1732, 'E', 1734-1736, 'T', 1738-1762, 'D', 1764-1848 <Y>  
R: Yamauchi-Takahara, K., Soie, M.U.; Liew, J.; Ing, D.; Liew, C.C.  
Proc. Natl. Acad. Sci. U.S.A. 86, 7416-7417, 1989

A: Reference number: A94226  
A: Contents: extratum  
A: Accession: B33835

A: Molecule type: DNA  
A: Residues: 1-87, 'Q', 89-177; 1551-1732, 'E', 1734-1736, 'T', 1738-1762, 'D', 1764-1939 <YA>  
R: Saez, L.V.; Gianola, K.M.; McCallally, E.M.; Feghalli, R.; Eddy, R.; Shows, T.B.; Leitwand  
Nucleic Acids Res. 15, 5443-5459, 1987

A: Title: Human cardiac myosin heavy chain genes and their linkage in the genome.  
A: Reference number: A93699; MUID:87260010, PMID:3037493

A: Accession: B27858  
A: Molecule type: DNA  
A: Residues: 1-3, 'S', 'S'-10, 'T', 12, 14-67 <SAE>  
R: Kurabayashi, M.; Teuchmochi, H.; Komuro, I.; Takaku, F.; Yazaki, Y.  
J. Clin. Invest. 82, 524-531, 1988

A: Title: Molecular cloning and characterization of human cardiac alpha- and beta-form my  
human actin.

A: Reference number: A92770; MUID:88299163; PMID:2969919  
A: Accession: A28908  
A: Molecule type: mRNA  
A: Residues: 1407-1532, 'N', 1534-1539, 'W', 1541-1576, 'NV', 1579-1704, 'EO', 1707-1762, 'D', 1764-  
A: Cross-references: GB:M1664; NID:g189006; PIDN:AAA6344.1; PID:g386972  
A: Experimental source: fetal heart  
C: Genetics:

A: Gene: GDB:MYH6  
A: Cross-references: GDB:120214; OMIM:160710  
A: Map position: 14q11.2-14q13

C: Superfamily: myosin heavy chain; myosin motor domain homology  
C: Keywords: actin binding; ATP; cardiac muscle; coiled coil; heart; hydrolase; methylated  
F: 88-766/Domain: myosin motor domain homology <MOTOR>  
F: 178-185/Region: nucleotide-binding motif A (P-loop)  
F: 549-586/Region: actin binding #status predicted  
F: 657-679/Region: actin binding #status predicted  
F: 841-1939/Domain: coiled coil #status predicted <COI>  
F: 841-1281/Region: S2  
F: 1282-1939/Region: light meromyosin  
F: 129/Modified site: NG/N6/N6-crimethylation (Lys) #status predicted  
F: 184/Binding site: ATP N6/N6 #status predicted  
F: 697, 707/Active site: Cys #status predicted

Query Match 19.2%; Score 75.5; DB 1; Length 1939;  
Best Local Similarity 30.3%; Pred. No. 16;  
Matches 20; Conservative 13; Mismatches 32; Indels 1; Gaps 1;

Oy 4 LDPLQLKLVFKIRRYKSKRTSGSPVADSSSYQDELRELFLKLQMGNDMTFFPKF 63  
Db 1444 LDKQRNFDDKIILAWKKQYESQSESSQKEARSLSTELFKLNAY-EESLEHLTEFR 1502

Oy 64 EDPKFE 69  
Db 1503 ENKNLIQ 1508

RESULT 15  
148175  
myosin heavy chain alpha, cardiac muscle [similarity] - golden hamster  
C: Species: Mesocricetus auratus (golden hamster)  
C: Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
A: Accession: 148175; A23938  
R: Wang, R.; Soie, M.U.; Cukerman, E.; Liew, C.C.  
J. Mol. Cell. Cardiol. 26, 1155-1165, 1994  
A: Title: Characterization and nucleotide sequence of the cardiac alpha-myosin heavy chain  
A: Reference number: 148153; MUID:95115033; PMID:7815459  
A: Accession: 148175  
A: Status: Preliminary; translated from GB/EMBL/DDBJ  
A: Molecule type: DNA  
A: Residues: 1-1939 <RES>  
A: Cross-references: UNIPROT:P13539; GB:L15351; NID:g402373; PIDN:AA59701.1; PID:g402374  
R: Liew, C.C.; Jandresek, W.A.  
Proc. Natl. Acad. Sci. U.S.A. 83, 3175-3179, 1986

```
A>Title: Construction and characterization of the alpha-form of a cardiac myosin heavy
A.Reference number: A23938; MUID:86205859; PMID:3458174
A.Accession: A23938
A.Molecule type: mRNA
A.Residues: 1630-1843, 'R',1845-1878, 'T',1880-1927,'N',1929-1932,1934-1939 <LIE>
C.Geneatics:
C.AIntrons: 67/3; 115/3; 168/1; 177/2; 214/3; 245/3; 267/1; 300/1; 334/3; 381/1; 470/3;,
392/2; 1453/3; 1509/1; 1550/3; 1653/3; 1721/3; 1763/3; 1855/3; 1887/3; 1932/3
C.Superfamily: myosin heavy chain; myosin motor domain homology
C.Keywords: ATP; cardiac muscle; heart; nucleotide binding; P-loop
F:88-768/Domain: myosin motor domain homology <MMOT>
F:178-185/Region: nucleotide-binding motif A (P-loop)

Query Match          19.2%; Score 75..5; DB 2; Length 1939;
Best Local Similarity 30.3%; Pred. No. 16;
Matches    20; Conservative   13; Mismatches     32; Indels      1; Gaps       1;
```

QY           4 LDPIDKLFDVDRKREYSKRQTSGSPVDASSSEIYOQLERELFKLIKONFGNDMNTFFPKKF 63  
       |||:::||::||::||::||::||::||  
Db      1444 LDKKRNPFDKILAEWKOKYEESQSLESSQKEARSLSTFLFKLNAY-BESELHLTEFPKR 1502

OY          64 EDPKFE 69  
       ||:  
Db      1503 ENKNLQ 1508

Search completed: April 4, 2005, 19:24:22  
Job time : 27 secs

Search completed: April 4, 2005, 19:24:22  
Job time : 27 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2005, 18:55:50 ; Search time 94.5 Seconds  
(without alignments)  
266.678 Million cell updates/sec

Title: US-09-831-951A-1  
Perfect score: 394  
Sequence: 1 NKEIDPIQKLFVDKIRREYKS.....TFPTKPEDPKFVLEKPKQA 76

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1413372 seqs, 311592847 residues

Total number of hits satisfying chosen parameters: 1413372

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	392	99.5	84	US-10-106-698-6594	Sequence 6594, Ap
2	392	99.5	108	US-10-131-410-119	Sequence 119, App
3	392	99.5	108	US-10-408-765A-1253	Sequence 1253, Ap
4	332	84.3	108	US-10-926-543-77	Sequence 77, Appl
5	112	28.4	56	US-09-864-761-33712	Sequence 33712, A
6	112	28.4	56	US-09-864-761-34125	Sequence 34125, A
7	79.5	20.2	1940	US-09-738-630-99	Sequence 99, Appl
8	79.5	20.2	1940	US-10-408-765A-1175	Sequence 1175, Ap
9	75.5	19.2	1137	US-10-336-472-16	Sequence 16, Appl
10	75.5	19.2	1185	US-10-340-797-3350	Sequence 3350, Ap
11	75.5	19.2	1859	US-10-336-472-20	Sequence 20, Appl
12	75.5	19.2	1935	US-10-336-472-22	Sequence 22, Appl
13	71	18.0	202	US-10-264-237-1904	Sequence 1904, Ap

14	71	18.0	567	15	US-10-108-260A-4464	Sequence 4464, Ap
15	70.5	17.9	892	14	US-10-205-219-102	Sequence 102, App
16	69.5	17.6	541	15	US-10-282-122A-77851	Sequence 77851, A
17	69.5	17.6	750	15	US-10-369-493-20521	Sequence 20521, A
18	69.5	17.6	1939	16	US-10-408-765A-2188	Sequence 2188, Ap
19	68.5	17.4	285	15	US-10-094-749-1655	Sequence 1655, Ap
20	68.5	17.4	332	15	US-10-108-260A-4506	Sequence 4506, Ap
21	67.5	17.1	744	15	US-10-369-493-7908	Sequence 7908, Ap
22	67.5	17.1	1941	16	US-10-437-963-171741	Sequence 171741, A
23	67.5	17.1	1983	15	US-10-282-122A-43976	Sequence 43976, A
24	67.5	17.1	2076	9	US-09-815-242-5815	Sequence 5815, Ap
25	67.5	17.1	2186	9	US-09-815-242-12913	Sequence 12913, A
26	67.5	17.1	2186	17	US-10-470-048B-336	Sequence 336, App
27	67	17.0	479	15	US-10-282-122A-72411	Sequence 72411, A
28	66.5	16.9	1184	15	US-10-282-122A-53254	Sequence 53254, A
29	66	16.8	711	9	US-09-748-875-3	Sequence 3, Appl
30	66	16.8	711	10	US-09-298-523B-3	Sequence 3, Appl
31	66	16.8	6641	15	US-10-282-122A-70580	Sequence 70580, A
32	66	16.8	10203	16	US-10-661-809-23	Sequence 23, Appl
33	65.5	16.6	507	15	US-10-282-122A-60276	Sequence 60276, A
34	65.5	16.6	740	15	US-10-369-493-21798	Sequence 21798, A
35	65.5	16.6	822	15	US-10-369-493-11834	Sequence 11834, A
36	65.5	16.6	2189	14	US-10-172-502-2	Sequence 2, Appl
37	65	16.5	651	15	US-10-614-076-52	Sequence 52, Appl
38	65	16.5	652	14	US-10-232-665-2	Sequence 2, Appl
39	65	16.5	652	14	US-10-232-665-4	Sequence 4, Appl
40	65	16.5	652	14	US-10-232-665-6	Sequence 6, Appl
41	65	16.5	652	15	US-10-614-076-4	Sequence 4, Appl
42	65	16.5	652	15	US-10-614-076-6	Sequence 6, Appl
43	65	16.5	652	15	US-10-614-076-8	Sequence 8, Appl
44	65	16.5	652	15	US-10-614-076-10	Sequence 10, Appl
45	65	16.5	652	15	US-10-614-076-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1  
US-10-106-698-6594  
; Sequence 6594, Application US/10106698  
; Publication No. US20030109690A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OR INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide  
; FILE REFERENCE: P4005P1  
; CURRENT APPLICATION NUMBER: US/10/106,698  
; PRIOR FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: PCT/US00/26524  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: US 60/163,280  
; PRIOR FILING DATE: 1999-11-03  
; NUMBER OF SEQ ID NOS: 8564  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 6594  
; LENGTH: 84  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-106-698-6594  
Query Match 99.5%; Score 392; DB 14; Length 84;  
Best Local Similarity 98.7%; Pred. No. 4.8e-38;  
Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
OY 1 NKEIDPIQKLFVDKIRREYKSGPVSASSYQOELRELFKLMKMGNAADMTPT 60  
DB 9 NKEIDPIQKLFVDKIRREYKSGPVSASSYQOELRELFKLMKMGNAADMTPT 68  
OY 61 FKPEDPKFVLEKPKQA 76  
DB 69 FKPEDPKFVLEKPKQA 84

```
RESULT 2
US-10-131-410-119
; Sequence 119, Application US/10131410
; Publication No. US20030235915A1
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM TISSUE OF BREAST
; FILE REFERENCE: SCH-1763
; CURRENT APPLICATION NUMBER: US/10/131,410
; PRIOR FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 09/646,673
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: PCT/DE99/00908
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 119
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-410-119
```

```
Query Match          99.5%; Score 392; DB 15; Length 108;
Best Local Similarity 98.7%; Pred. No. 6,5e-38;
Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 NKEIDPIQKLFVDKIREYKSKRQTSQGPVDASSEYQOELERELFKLKOMFGNADMTPT 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 33 NKEIDPIQKLFVDKIREYKSKRQTSQGPVDASSEYQOELERELFKLKOMFGNADMTPT 92
QY 61 FKFEDEPKFEVLEKPOA 76
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 93 FKFEDEPKFEVLEKPOA 108
```

```
RESULT 3
US-10-408-765A-1253
; Sequence 1253, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Boia D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary W.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088,465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1253
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1253
```

```
Query Match          99.5%; Score 392; DB 16; Length 108;
Best Local Similarity 98.7%; Pred. No. 6,5e-38;
Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 NKEIDPIQKLFVDKIREYKSKRQTSQGPVDASSEYQOELERELFKLKOMFGNADMTPT 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
DB 33 NKEIDPIQKLFVDKIREYKSKRQTSQGPVDASSEYQOELERELFKLKOMFGNADMTPT 92
QY 61 FKFEDEPKFEVLEKPOA 76
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 93 FKFEDEPKFEVLEKPOA 108
```

```
RESULT 4
US-10-926-543-77
; Sequence 77, Application US/10926543
; Publication No. US20050048589A1
; GENERAL INFORMATION:
; APPLICANT: Jendoubi, Moncef
; TITLE OF INVENTION: LUNG CANCER SPECIFIC GENE PRODUCTS: THEIR CODING SEQUENCE, THEIR
; TITLE OF INVENTION: ANTIBODIES AND THEIR USE IN DIAGNOSTIC, THERAPEUTIC AND DISEASE
; TITLE OF INVENTION: MANAGEMENT OF LUNG CANCER
; FILE REFERENCE: 705403,4004
; CURRENT APPLICATION NUMBER: US/10/926,543
; CURRENT FILING DATE: 2004-08-25
; PRIOR APPLICATION NUMBER: US 60/497,790
; PRIOR FILING DATE: 2003-08-25
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 77
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-926-543-77
```

```
Query Match          84.3%; Score 332; DB 17; Length 108;
Best Local Similarity 77.6%; Pred. No. 6,6e-31;
Matches 59; Conservative 11; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 NKEIDPIQKLFVDKIREYKSKRQTSQGPVDASSEYQOELERELFKLKOMFGNADMTPT 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 33 NKEIDPIQKLFVDKIREYKSKRQTSQGPVDIGPEYQOEDIDRELYKXKMGKMGMDTPT 92
QY 61 FKFEDEPKFEVLEKPOA 76
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 93 FKFEDEPKFEVIDKPOA 108
```

```
RESULT 5
US-09-864-761-33712
; Sequence 33712, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
```



```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33712
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000087.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
; OTHER INFORMATION: SWISSPROT HIT: P18859, EVALUE 5.00e-24
; OTHER INFORMATION: EST_HUMAN HIT: AA128761.1, EVALUE 5.00e-24
US-09-864-761-33712

Query Match      28.4%; Score 112; DB 9; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 NKEIDPIOKLFVDKIREYKSR 22
      |||
Db      35 NKEIDPIOKLFVDKIREYKSR 56

RESULT 6
US-09-864-761-34125
; Sequence 34125, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aegmca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
```

```
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34125
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000139.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EST_HUMAN HIT: AA128761.1, EVALUE 5.00e-24
; OTHER INFORMATION: SWISSPROT HIT: P18859, EVALUE 5.00e-24
US-09-864-761-34125

Query Match      28.4%; Score 112; DB 9; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 NKEIDPIOKLFVDKIREYKSR 22
      |||
Db      35 NKEIDPIOKLFVDKIREYKSR 56

RESULT 7
US-09-738-630-99
; Sequence 99, Application US/09738630
; Publication No. US20030166213A1
; GENERAL INFORMATION:
; APPLICANT: Greenpan, Ralph J.
; APPLICANT: Shaw, Paul J.
; TITLE OF INVENTION: Methods For Identifying Compounds That
; FILE REFERENCE: P-NI 3906
; TITLE OF INVENTION: Modulate Disorders Related To Nitric Oxide/cGMP-Dependent
; FILE REFERENCE: P-NI 3906
; CURRENT APPLICATION NUMBER: US/09/738,630
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
```

SEQ ID NO 99  
LENGTH: 1940  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-738-630-99

Query Match 20.2%; Score 79.5; DB 10; Length 1940;  
Best Local Similarity 33.3%; Pred. No. 7;  
Matches 22; Conservative 13; Mismatches 30; Indels 1; Gaps 1;

QY 4 LDPQLKLVFKIRKYSKRGTSQSGPVDAASSEYQQLERELFKLQKMGADMTPTPKF 63  
DB 1443 LDKKQRPNDKVLAEWKTKCESQAELESLKESRSLSTELFKLKNAYEBA-LDQLETVKR 1501

QY 64 EDPKFE 69  
DB 1502 ENKNLE 1507

RESULT 8  
US-10-408-765A-1175  
Sequence 1175, Application US/10408765A  
Publication No. US20040101874A1  
GENERAL INFORMATION:  
APPLICANT: Ghosh, Soumitra S.  
APPLICANT: Fany, Eoin D.  
APPLICANT: Zhang, Bing  
APPLICANT: Gibson, Bradford W.  
APPLICANT: Taylor, Steven W.  
APPLICANT: Glenn, Gary W.  
APPLICANT: Warnock, Dale E.  
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
FILE REFERENCE: 660088.465  
CURRENT APPLICATION NUMBER: US/10/408,765A  
CURRENT FILING DATE: 2003-04-04  
NUMBER OF SEQ ID NOS: 3077  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1175  
LENGTH: 1940  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-408-765A-1175

Query Match 20.2%; Score 79.5; DB 16; Length 1940;  
Best Local Similarity 33.3%; Pred. No. 7;  
Matches 22; Conservative 13; Mismatches 30; Indels 1; Gaps 1;

QY 4 LDPQLKLVFKIRKYSKRGTSQSGPVDAASSEYQQLERELFKLQKMGADMTPTPKF 63  
DB 1443 LDKKQRPNDKVLAEWKTKCESQAELESLKESRSLSTELFKLKNAYEBA-LDQLETVKR 1501

QY 64 EDPKFE 69  
DB 1502 ENKNLE 1507

RESULT 9  
US-10-336-472-16  
Sequence 16, Application US/10336472  
Publication No. US20040043929A1  
GENERAL INFORMATION:  
APPLICANT: Anderson, David W.  
APPLICANT: Baillinger, Robert A.  
APPLICANT: Baumgartner, Jason C.  
APPLICANT: Burgess, Catherine E.  
APPLICANT: Casman, Seacie J.  
APPLICANT: Chant, John S.  
APPLICANT: Berghs, Constance  
APPLICANT: Gangoli, Beha A.  
APPLICANT: Edinger, Shlomit R.  
APPLICANT: Ellerman, Karen  
APPLICANT: Furtak, Katarzyna

Query Match 19.2%; Score 75.5; DB 15; Length 1137;  
Best Local Similarity 30.3%; Pred. No. 11;  
Matches 20; Conservative 13; Mismatches 32; Indels 1; Gaps 1;

QY 4 LDPQLKLVFKIRKYSKRGTSQSGPVDAASSEYQQLERELFKLQKMGADMTPTPKF 63  
DB 644 LDKKQRPNDKVLAEWKTKCESQAELESLKESRSLSTELFKLKNAY-EESLHLETVKR 702

QY 64 EDPKFE 69  
DB 703 ENKNLE 708

APPLICANT: Gerlach, Valerie  
APPLICANT: Gilbert, Jennifer A.  
APPLICANT: Gunther, Erik  
APPLICANT: Gorman, Linda  
APPLICANT: Guo, Xiaojia Sasha  
APPLICANT: Li, Weizhen  
APPLICANT: Liu, Xiaohong  
APPLICANT: Miller, Charles E.  
APPLICANT: Padigar, Muralidhara  
APPLICANT: Patutajan, Meera  
APPLICANT: Rastelli, Luca  
APPLICANT: MacDougall, John R.  
APPLICANT: Mishra, Vishnu  
APPLICANT: Pena, Carol E.A.  
APPLICANT: Spaderna, Steven K.  
APPLICANT: Shinkets, Richard A.  
APPLICANT: Smithson, Glenda  
APPLICANT: Spytek, Kimberly A.  
APPLICANT: Stone, David J.  
APPLICANT: Shenoy, Suresh G.  
APPLICANT: Ort, Tatiana  
APPLICANT: Taupier Jr, Raymond J.  
APPLICANT: Tchernyev, Velizar T.  
APPLICANT: Vernet, Corine A.M.  
APPLICANT: Wolenc, Adam R.  
APPLICANT: Zerhusen, Bryan D.  
APPLICANT: Zhong, Mei  
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: 21402-533C  
CURRENT APPLICATION NUMBER: US/10/336,472  
CURRENT FILING DATE: 2003-01-03  
PRIOR APPLICATION NUMBER: 09/746,491  
PRIOR FILING DATE: 2000-12-20  
PRIOR APPLICATION NUMBER: 10/005,041  
PRIOR FILING DATE: 2001-12-04  
PRIOR APPLICATION NUMBER: 10/023,681  
PRIOR FILING DATE: 2001-12-18  
PRIOR APPLICATION NUMBER: 10/024,212  
PRIOR FILING DATE: 2001-12-18  
PRIOR APPLICATION NUMBER: 10/055,569  
PRIOR FILING DATE: 2001-10-26  
PRIOR APPLICATION NUMBER: 10/080,334  
PRIOR FILING DATE: 2002-02-21  
PRIOR APPLICATION NUMBER: 10/092,900  
PRIOR FILING DATE: 2002-03-07  
PRIOR APPLICATION NUMBER: 10/136,826  
PRIOR FILING DATE: 2002-05-01  
PRIOR APPLICATION NUMBER: 10/236,417  
PRIOR FILING DATE: 2002-09-06  
PRIOR APPLICATION NUMBER: 60/345,092  
PRIOR FILING DATE: 2002-01-04  
Remaining Prior Application data removed - See file wrapper or PAM.  
NUMBER OF SEQ ID NOS: 230  
SOFTWARE: CuraSegList version 0.1  
SEQ ID NO 16  
LENGTH: 1137  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-336-472-16

```
RESULT 10
US-10-320-797-3350
; Sequence 3350, Application US/10320797
; Publication No. US20040014955A1
; GENERAL INFORMATION:
; APPLICANT: Eroshkin, Alexey M.
; APPLICANT: Zamudio, Carlos
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 10182-021-999
; CURRENT APPLICATION NUMBER: US/10/320,797
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,261
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 3361
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3350
; LENGTH: 1185
; TYPE: PRT
; ORGANISM: Cryptococcus neoformans
US-10-320-797-3350

Query Match          19.2%; Score 75.5; DB 15; Length 1185;
Best Local Similarity 39.1%; Pred. No. 11;
Matches 18; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

Qy      2 KELDPIQKLFVDKIRKYSKRQTSGGPVDASSSEYQOELERELFKIK 47
Db      278 KECHGMEK-ELBEIRKKKEQAKGKIQGLTEAVNNLERELVKIK 322

RESULT 11
US-10-336-472-20
; Sequence 20, Application US/10336472
; Publication No. US20040043929A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.
; APPLICANT: Ballinger, Robert A.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Chant, John S.
; APPLICANT: Berghs, Constance
; APPLICANT: Gangolli, Esna A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gilbert, Jennifer A.
; APPLICANT: Gunther, Erik
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Ji, Weizhen
; APPLICANT: Liu, Xiaohong
; APPLICANT: Miller, Charles E.
; APPLICANT: Miller, Isabelle
; APPLICANT: Padigaru, Muralihara
; APPLICANT: Patturajan, Meera
; APPLICANT: Raetelli, Luca
; APPLICANT: MacDougall, John R.
; APPLICANT: Mishra, Vishnu
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Shinkens, Richard A.
; APPLICANT: Smithson, Glenda
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Stone, David J.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Ort, Tatiana
; APPLICANT: Taupier Jr, Raymond J.
```

```
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Zerkusen, Bryan D.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-533C
; CURRENT APPLICATION NUMBER: US/10/336,472
; CURRENT FILING DATE: 2003-01-03
; PRIOR APPLICATION NUMBER: 09/746,491
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 10/005,041
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 10/023,681
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 10/024,212
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 10/055,569
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: 10/080,334
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 10/092,900
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 10/136,826
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 10/236,417
; PRIOR FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 60/345,092
; PRIOR FILING DATE: 2002-01-04
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 20
; LENGTH: 1859
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-336-472-20

Query Match          19.2%; Score 75.5; DB 15; Length 1859;
Best Local Similarity 30.3%; Pred. No. 19;
Matches 20; Conservative 13; Mismatches 32; Indels 1; Gaps 1;

Qy      4 LDPIQKLFVDKIRKYSKRQTSGGPVDASSSEYQOELERELFKIKQMGNADMTFPTPKF 63
Db      1442 LDKQKRNFDKILAEWKQKREBSQSELSQKARSISTELFKLNAT-EESLEHLETPKR 1500

Qy      64 EDPKFE 69
Db      1501 ENKMLQ 1506

RESULT 12
US-10-336-472-22
; Sequence 22, Application US/10336472
; Publication No. US20040043929A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.
; APPLICANT: Ballinger, Robert A.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Chant, John S.
; APPLICANT: Berghs, Constance
; APPLICANT: Gangolli, Esna A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gilbert, Jennifer A.
; APPLICANT: Gunther, Erik
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Ji, Weizhen
```



## RESULT 15

US-10-205-219-102

; Sequence 102, Application US/10205219

; Publication No. US20030138803A1

; GENERAL INFORMATION:

; APPLICANT: Warner-Lambert Company

; APPLICANT: Lee, Kevin

; APPLICANT: Dixon, Alistair

; APPLICANT: Brookbank, Robert

; APPLICANT: Pincock, Robert

; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain

; FILE REFERENCE: WL-A-018200

; CURRENT APPLICATION NUMBER: US/10/205,219

; CURRENT FILING DATE: 2002-07-24

; PRIOR APPLICATION NUMBER: GB 0118354.0

; PRIOR FILING DATE: 2001-07-27

; NUMBER OF SEQ ID NOS: 197

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 102

; LENGTH: 892

; TYPE: PRT

; ORGANISM: Rattus norvegicus

; FEATURE:

; OTHER INFORMATION: Myosin heavy chain

US-10-205-219-102

Query Match

Best Local Similarity 17.9%; Score 70.5; DB 14; Length 892;

Matches 18; Conservative 15; Mismatches 32; Indels 1; Gaps 1;

OY 4 LDPQLFVDFKIREYKSKRTSGPVDASSEYQELERELPFLKQMFQGNADWNTPTPKF 63  
DB 399 LDKQRNFDKILAEWKQKYEETHTAELEASQKESRSLSTELFKIKNAV-EESSLDQLETLKR 457  
OY 64 EDPKFE 69  
DB 458 ENKNLQ 463

Search completed: April 4, 2005, 19:15:02  
Job time : 95.5 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2005, 18:46:29 ; Search time 30 Seconds  
(without alignments)  
189.111 Million cell updates/sec

Title: US-09-831-951A-1  
Perfect score: 394  
Sequence: 1 NKEIDPIQKLVFKIREYKS.....TFPTKFEVLEKRPQA 76

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	392	99.5	108	4	US-09-513-999C-8086
2	392	99.5	108	4	US-09-949-016-11049
3	387	98.2	108	2	US-08-828-239-3
4	387	98.2	108	3	US-09-205-679-3
5	383	97.2	106	4	US-09-513-999C-5967
6	346	87.8	108	2	US-08-828-239-4
7	346	87.8	108	3	US-09-205-679-4
8	325	82.5	114	2	US-08-828-239-1
9	325	82.5	114	3	US-09-205-679-1
10	317	80.5	108	2	US-08-828-239-5
11	317	80.5	108	3	US-09-205-679-5
12	185	47.0	69	4	US-09-513-999C-5968
13	79.5	20.2	1940	4	US-09-538-092-901
14	79.5	20.2	1963	4	US-09-949-016-8888
15	75.5	19.2	1886	4	US-08-938-105-3
16	75.5	19.2	1935	4	US-09-538-092-916
17	75.5	19.2	1939	3	US-09-310-187A-1
18	75.5	19.2	1939	4	US-09-538-092-917
19	75.5	19.2	1944	4	US-09-949-016-10929
20	69.5	17.6	1937	4	US-09-538-092-918
21	69.5	17.6	1939	4	US-09-538-092-919
22	69.5	17.6	1939	4	US-09-949-016-6925
23	69.5	17.6	1939	4	US-09-949-016-11104
24	69.5	17.6	1942	4	US-09-949-016-8135
25	68.5	17.4	215	4	US-09-949-016-7173
26	68.5	17.4	217	1	US-08-390-858B-9
27	68	17.3	284	1	US-08-390-858B-36

28	67.5	17.1	134	4	US-09-248-796A-17771	Sequence 17771, A
29	67.5	17.1	555	3	US-09-813-872-4	Sequence 4, Appl1
30	66	16.8	706	4	US-09-538-092-1266	Sequence 1266, Ap
31	66	16.8	1095	4	US-09-710-279-3154	Sequence 3154, Ap
32	66	16.8	10182	3	US-09-134-001C-3159	Sequence 3159, Ap
33	65.5	16.6	206	4	US-09-949-016-6411	Sequence 6411, Ap
34	65.5	16.6	520	4	US-09-489-039A-11079	Sequence 11079, A
35	65	16.5	551	3	US-08-996-441B-52	Sequence 52, Appl1
36	65	16.5	651	3	US-08-993-170A-52	Sequence 52, Appl1
37	65	16.5	651	3	US-08-993-170A-52	Sequence 52, Appl1
38	65	16.5	651	3	US-08-993-170A-52	Sequence 52, Appl1
39	65	16.5	651	4	US-09-427-770-52	Sequence 52, Appl1
40	65	16.5	651	4	US-09-427-769-52	Sequence 52, Appl1
41	65	16.5	652	3	US-08-996-441B-6	Sequence 4, Appl1
42	65	16.5	652	3	US-08-996-441B-6	Sequence 6, Appl1
43	65	16.5	652	3	US-08-996-441B-8	Sequence 8, Appl1
44	65	16.5	652	3	US-08-996-441B-10	Sequence 10, Appl1
45	65	16.5	652	3	US-08-996-441B-12	Sequence 12, Appl1

## ALIGNMENTS

```

RESULT 1
US-09-513-999C-8086
; Sequence 8086, Application US/0951399C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59. US2. REG
; CURRENT APPLICATION NUMBER: US/09/513.999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 8086
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-8086

Query Match          99.5%; Score 392; DB 4; Length 108;
Best Local Similarity 98.7%; Pred. No. 1.4e-39;
Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 1 NKEIDPIQKLVFKIREYKSGPVDASSYQOELRELFKTKOMFGNADMTFFPT 60
    |||
Db 33 NKEIDPIQKLVFKIREYKSKRQTSGGPVDASSYQOELRELFKTKOMFGNADMTFFPT 92
    |||

Cy 61 FKFEVLEKRPQA 76
    |||
Db 93 FKFEVLEKRPQA 108
    |||

RESULT 2
US-09-949-016-11049
; Sequence 11049, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949.016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

```

;; PRIOR FILING DATE: 2000-10-03  
;; PRIOR APPLICATION NUMBER: 60/231,498  
;; PRIOR FILING DATE: 2000-09-08  
;; NUMBER OF SEQ ID NOS: 207012  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 11049  
;; LENGTH: 108  
;; TYPE: PR1  
;; ORGANISM: Human  
US-09-949-016-11049

Query Match 99.5%; Score 392; DB 4; Length 108;  
Best Local Similarity 98.7%; Pred. No. 1,4e-39;  
Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKELDPIQKLFVNDKIREYKSKRQTSQGPVDASSEYOOELERELFKLKQMGNDMNTFPT 60  
DB 33 NKELDPIQKLFVNDKIREYKSKRQTSQGPVDASSEYOOELERELFKLKQMGNDMNTFPT 92  
QY 61 FKFEDEPKFEVLEKPOA 76  
DB 93 FKFEDEPKFEVLEKPOA 108

RESULT 3  
US-08-828-239-3  
Sequence 3, Application US/08828239  
Patent No. 5849527  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Shah, Puri  
TITLE OF INVENTION: NOVEL ATP SYNTHASE COUPLING FACTOR 6  
NUMBER OF SEQUENCES: 5  
CURRENT APPLICATION DATA:  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/828,239  
FILING DATE: Herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0260 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 108 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 183786  
US-08-828-239-3

Query Match 98.2%; Score 387; DB 2; Length 108;  
Best Local Similarity 97.4%; Pred. No. 5.7e-39;

Matches 74; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKELDPIQKLFVNDKIREYKSKRQTSQGPVDASSEYOOELERELFKLKQMGNDMNTFPT 60  
DB 33 NKELDPIQKLFVNDKIREYKSKRQTSQGPVDASSEYOOELERELFKLKQMGNDMNTFPT 92  
QY 61 FKFEDEPKFEVLEKPOA 76  
DB 93 FKFEDEPKFEVLEKPOA 108

RESULT 4  
US-09-205-679-3  
Sequence 3, Application US/09205679  
Patent No. 6048718  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Shah, Puri  
TITLE OF INVENTION: NOVEL ATP SYNTHASE COUPLING FACTOR 6  
NUMBER OF SEQUENCES: 5  
CURRENT APPLICATION DATA:  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/205,679  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/828,239  
FILING DATE: 03/31/1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0260 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 108 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 183786  
US-09-205-679-3

Query Match 98.2%; Score 387; DB 3; Length 108;  
Best Local Similarity 97.4%; Pred. No. 5.7e-39;  
Matches 74; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKELDPIQKLFVNDKIREYKSKRQTSQGPVDASSEYOOELERELFKLKQMGNDMNTFPT 60  
DB 33 NKELDPIQKLFVNDKIREYKSKRQTSQGPVDASSEYOOELERELFKLKQMGNDMNTFPT 92  
QY 61 FKFEDEPKFEVLEKPOA 76  
DB 93 FKFEDEPKFEVLEKPOA 108

RESULT 5  
US-09-513-999C-5967



```
; Sequence 5967, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5967
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-5967

Query Match          97.2%; Score 383; DB 4; Length 106;
Best Local Similarity 98.6%; Pred. No. 1.7e-38;
Matches 73; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKEIDPIQKLFVDKIREYKSKRQTSGGPVDASSSYQQLERELFKLKQMPGNADMTPT 60
DB 33 NKEIDPIQKLFVDKIREYKSKRQTSGGPVDASSSYQQLERELFKLKQMPGNADMTPT 92

QY 61 FKPEDPKFEVLEKRP 74
DB 93 FKPEDPKFEVLEKRP 106

RESULT 6
US-08-828-239-4
; Sequence 4, Application US/08828239
; Patent No. 5849527
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: NOVEL ATP SYNTHASE COUPLING FACTOR 6
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/828,239
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0260 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 108 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Genbank
; CLONE: 625323
US-09-205-679-4
```

```
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Genbank
; CLONE: 625323
US-08-828-239-4

Query Match          87.8%; Score 346; DB 2; Length 108;
Best Local Similarity 82.9%; Pred. No. 4.7e-34;
Matches 63; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 NKEIDPIQKLFVDKIREYKSKRQTSGGPVDASSSYQQLERELFKLKQMPGNADMTPT 60
DB 33 NKEIDPIQKLFVDKIREYKSKRQTSGGPVDASSSYQQLERELFKLKQMPGNADMTPT 92

QY 61 FKPEDPKFEVLEKPOA 76
DB 93 FTPEPKFEVLEKPOS 108

RESULT 7
US-09-205-679-4
; Sequence 4, Application US/09205679
; Patent No. 6048718
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: NOVEL ATP SYNTHASE COUPLING FACTOR 6
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/205,679
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/828,239
; FILING DATE: 03/31/1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0260 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 108 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Genbank
; CLONE: 625323
US-09-205-679-4

Query Match          87.8%; Score 346; DB 3; Length 108;
Best Local Similarity 82.9%; Pred. No. 4.7e-34;
Matches 63; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 NKEIDPIQKLFVDKIREYKSKRQTSGGPVDASSSYQQLERELFKLKQMPGNADMTPT 60
```

Db 33 NKEIDPVOKLFVDKIREYTRKQTSQGVADGPEYQODLDRLEFKLKQMGKADMTFBN 92  
QY 61 FKFEDEPKFEVLEKPOA 76  
Db 93 FTFEDEPKFEVLEKPOS 108

## RESULT 8

US-08-828-239-1

; Sequence 1, Application US/08828239  
; Patent No. 5849527  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: NOVEL ATP SYNTHASE COUPLING FACTOR 6  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/828,239  
; FILING DATE: Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0260 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 114 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: Consensus  
; CLONE: 2496341  
; US-08-828-239-1

Query Match 82.5%; Score 325; DB 2; Length 114;  
Best Local Similarity 98.4%; Pred. No. 1.6e-31;  
Matches 62; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 14 KIREYKSKRQTSQGVADASSYQOELERELFKLKQMGFNADMTFPTFKFEDEPKFEVLEK 73  
Db 52 KIREYKSKRQTSQGVADASSYQOELERELFKLKQMGFNADMTFPTFKFEDEPKFEVLEK 111  
QY 74 POA 76  
Db 112 POA 114

RESULT 9  
US-09-205-679-1  
; Sequence 1, Application US/09205679  
; Patent No. 6048718  
; GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: NOVEL ATP SYNTHASE COUPLING FACTOR 6  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/205,679  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/828,239  
; FILING DATE: 03/31/1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0260 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 114 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: Consensus  
; CLONE: 2496341  
; US-09-205-679-1

Query Match 82.5%; Score 325; DB 3; Length 114;  
Best Local Similarity 98.4%; Pred. No. 1.6e-31;  
Matches 62; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 14 KIREYKSKRQTSQGVADASSYQOELERELFKLKQMGFNADMTFPTFKFEDEPKFEVLEK 73  
Db 52 KIREYKSKRQTSQGVADASSYQOELERELFKLKQMGFNADMTFPTFKFEDEPKFEVLEK 111  
QY 74 POA 76  
Db 112 POA 114

RESULT 10  
US-08-828-239-5  
; Sequence 5, Application US/08828239  
; Patent No. 5849527  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: NOVEL ATP SYNTHASE COUPLING FACTOR 6  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette

RESULT 11  
 US-09-205-679-5  
 Sequence 5, Application US/09205679  
 Patent No. 6048718  
 GENERAL INFORMATION:  
 APPLICANT: Hillman, Jennifer L.  
 APPLICANT: Shah, Purvi  
 TITLE OF INVENTION: NOVEL ATP SYNTHASE COUPLING FACTOR 6  
 NUMBER OF SEQUENCES: 5  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 STREET: 3174 Porter Drive  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/205,679  
 FILING DATE: Herewith  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/828,239  
 FILING DATE: 03/31/1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Billings, Lucy J.  
 REGISTRATION NUMBER: 36,749

```

RESULT 12
US-09-513-999C-5968
; Sequence 5968, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J. B.
; APPLICANT: Duclerc, A.
; APPLICANT: Giordano, J. Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513.999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5968
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 65
; OTHER INFORMATION: Xaa-Pro or Ser
US-09-513-999C-5968

Query Matchch          47.0%; Score 185; DB 4; Length 69;
Best Local Similarity  97.3%; Pred. No. 5.3e-15;
Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 NKELDPIQKLFVDKIRFYKSKRQTSQGPVDAASRYQQ 37
        |||||||
Db       33 NKELDPIQKLFVDKIRFYKSKRQTSQGPVDAASRYQQ 69
        |||||||

RESULT 13
US-09-538-992-901
; Sequence 901, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Gioc, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same

```

FILE REFERENCE: 15966-542  
CURRENT APPLICATION NUMBER: US/09/538,092  
CURRENT FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/127,352  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/178,965  
PRIOR FILING DATE: 2000-02-01  
NUMBER OF SEQ ID NOS: 1387  
SOFTWARE: Curataseqformatler Version 0.9  
SEQ ID NO 901  
LENGTH: 1940  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (0)...(0)  
OTHER INFORMATION: Polypeptide Accession Number P11055  
US-09-538-092-901

Query Match 20.2%; Score 79.5; DB 4; Length 1940;  
Best Local Similarity 33.3%; Pred. No. 1.4;  
Matches 22; Conservative 13; Mismatches 30; Indels 1; Gaps 1;

QY 4 LDPLOKLFVNDKIRKRYKSKRQTSGGPVDASSEYQOELERELFKLKQMGNDMTPTPFKF 63  
DB 1443 LDKKQRPNDKVLAEWKTKCESQAELESLKESLSLSTELFKLKNAYEEA-IDQLETVYKR 1501

QY 64 EDPKFE 69  
DB 1502 ENKNLE 1507

RESULT 14  
US-09-949-016-8888  
Sequence 8888, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: C0001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8888  
LENGTH: 1963  
TYPE: PRT  
ORGANISM: Human  
US-09-949-016-8888

Query Match 20.2%; Score 79.5; DB 4; Length 1963;  
Best Local Similarity 33.3%; Pred. No. 1.4;  
Matches 22; Conservative 13; Mismatches 30; Indels 1; Gaps 1;

QY 4 LDPLOKLFVNDKIRKRYKSKRQTSGGPVDASSEYQOELERELFKLKQMGNDMTPTPFKF 63  
DB 1466 LDKKQRPNDKVLAEWKTKCESQAELESLKESLSLSTELFKLKNAYEEA-IDQLETVYKR 1524

QY 64 EDPKFE 69  
DB 1525 ENKNLE 1530

RESULT 15  
US-08-938-105-3  
Sequence 3, Application US/08938105

Patent No. 6353151  
GENERAL INFORMATION:  
APPLICANT: leinwand, Leslie A.  
APPLICANT: Vikstrom, Karen L.  
TITLE OF INVENTION: TRANSGENIC MODEL FOR HEART FAILURE  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross P.C.  
STREET: 1700 Lincoln St., Suite 3500  
CITY: Denver  
STATE: CO  
COUNTRY: U.S.A.  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/938,105  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Crook, Wanneil M.  
REGISTRATION NUMBER: 31,071  
REFERENCE/DOCKET NUMBER: 3595-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1886 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-938-105-3

Query Match 19.2%; Score 75.5; DB 3; Length 1886;  
Best Local Similarity 30.3%; Pred. No. 4.1;  
Matches 20; Conservative 13; Mismatches 32; Indels 1; Gaps 1;

QY 4 LDPLOKLFVNDKIRKRYKSKRQTSGGPVDASSEYQOELERELFKLKQMGNDMTPTPFKF 63  
DB 1391 LDKKQRPNDKVLAEWKTKCESQAELESLKESLSLSTELFKLKNAY-EESSHLETVYKR 1449

QY 64 EDPKFE 69  
DB 1450 ENKNLQ 1455

Search completed: April 4, 2005, 18:59:20  
Job time : 31 secs